Eigenvalue Bounds for Matrices

P.N. Khambule

January 2018

Eigenvalue Bounds for Matrices

by

Pretty Nombuyiselo Khambule

Submitted in fulfilment of the

academic requirements for the degree of

Master of Science

in the

School of

Mathematics, Statistics and Computer Science

University of KwaZulu-Natal

Durban

January 2018

As the candidate's supervisors, we have approved this thesis for submission.

Signed:............................ Name: Dr P.Singh............. Date:

Signed:............................ Name: Dr V.Singh............. Date:

Preface

The work described in this dissertation was carried out in the School of Mathematical Sciences, University of KwaZulu-Natal, Durban, from February 2017 to January 2018, under the supervision of Dr P. Singh and co-supervised by Dr V. Singh. This study represents original work by the author and has not otherwise been submitted in any form for any degree or diploma to any other tertiary institution. Where use has been made of the work of others, it is duly acknowledged in the text.

Signed:............................

Miss Pretty Nombuyiselo Khambule

FACULTY OF SCIENCE AND AGRICULTURE

DECLARATION 1 - PLAGIARISM

I, Pretty Nombuyiselo Khambule, declare that

1. The research reported in this thesis represents my own original research, except where due reference is made.

2. This thesis has not been submitted in the past for any degree at any other institution (academic or otherwise).

3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from the said persons.

4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from the said persons. Where other written sources have been quoted, then:

   a. Their words have been re-written but the general information attributed to them has been referenced.

   b. Where their exact words have been used, their writing has then been placed in italics, within quotation marks, and referenced.

5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the Bibliography.

Signed:..

Miss Pretty Nombuyiselo Khambule

In Memoriam

Thandazile May Madwe

(1969-2013)

Acknowledgement

Firstly, I would like to thank my supervisors, beginning with Dr. Pravin Singh, for all your support, patience and guidance which result in the completion of this thesis. Thank you for being one of those supervisors whom I was able to approach at any time, with any academic, artistic or social topic, and sort of rant and question. Your insights, passion, faith and kindness have proved invaluable to my education and to my life. To my co-supervisor Dr. Virath Singh, thank you for giving me the opportunity to do my masters degree under your mentorship. Thank you most of all, for your faith in me as student. This collaboration with the two of you has truly been one which I shall treasure always.

My special thanks also goes out to the following:

- To Professor Sunil Maharaj for financial assistance in the form of a bursary during the course of this study.
- My colleagues in the school of Mathematics, Statistics and Computer Science for their support and assistance.
- Members of staff in the School of Mathematics, Statistics and Computer Science for their on-going support and encouragement.
- The University of KwaZulu-Natal for granting me the opportunity to take on and complete this study.
- My dear friends whom have shared every moment of sadness and happiness with me.

- Lastly, my Ancestors for guiding me throughout this study. My grandmother, my pillar, thank you for your tireless support and your prayers. Also to my family for their support, as well as the sacrifices made on my behalf.

Notation

λ	Eigenvalue
$\|\cdot\|$	Norm
$\langle \cdot, \cdot \rangle$	Inner product
$\sigma(\cdot)$	Spectrum of operator $\cdot$
$(\cdot)^*$	Adjoint of operator $\cdot$
$(\cdot)^T$	transpose of operator $\cdot$
$tr(\cdot)$	Trace of operator $\cdot$
$\mathbb{C}$	Set of complex numbers
$\mathbb{R}$	Set of real numbers
δ_{ij}	Kronecker delta

Abstract

Eigenvalues are characteristic of linear operators. Once the spectrum of a matrix is known then its Jordan Canonical form can be determined which simplifies the understanding of the matrix. For large matrices and spectral analysis sometimes it is only necessary to know the eigenvalues of smallest and largest absolute values. Hence we consider various strategies of bounding the spectrum in the complex plane. Such bounds may be numerically improved by various algorithms. The minimal and maximal eigenvalues are crucial to determine the condition number of linear systems.

Contents

List of Figures

Chapter 1

Introduction

Eigenvalues are a special set of scalars associated with linear systems. They are also known as characteristic roots, characteristic values, proper values or spectral values. Historically, they arose in the study of quadratic forms and differential equations. Eigenvalues characterize important properties of linear transformations, being linked to invariant subspaces.

In the eighteenth century Euler studied the rotational motion of a rigid body and discovered the importance of the principal axes[6]. Lagrange discovered that the principal axes are the eigenvectors of the inertia matrix[9]. In the early nineteenth century, Cauchy used this work to classify the quadric surfaces and generalized it to arbitrary dimensions[9]. At the start of the 20th century, Hilbert studied the eigenvalues of integral operators by viewing the operators as matrices of infinite dimension[12].

In 1929 Von Mises published his work on the power method to compute the eigenvalues and eigenvectors of finite dimensional matrices[16]. Today the QR algorithm is one of the most efficient methods used to numerically compute eigenvalues and eigenvectors. It was independently discovered by France J.G.F[10] and Kublanovskaya V[13] in 1961.

Eigenvalues and eigenvectors are widely used in science and engineering. Civil engineers use eigenvalues to analyse and model physical systems arising in the design of bridges. The eigenvalues of the smallest magnitude represents the natural frequency of the bridge and is used to ensure stability of the structure[20].

In electrical engineering, the application of eigenvalues and eigenvectors is useful for decoupling three-phase systems through symmetrical component transformation. They are also used to determine the stability of electrical machines[23].

Eigenvalue analysis is also used in the design of car stereos systems, where it helps to reduce the vibration of the car due to music[1]. Eigenvalues and eigenvectors can also be used to test for cracks or deformities in a solid. Oil companies use eigenvalues analysis to explore land for oil. Since oil, dirt, and other substances give rise to linear systems which have different eigenvalues, hence this can be used to locate hidden oil reserves[1].

Claude Shannon used eigenvalues to determine the theoretical limit to how much information can be transmitted through a communication medium like your telephone line or through the air. This is done by calculating the eigenvectors and eigenvalues of the communication channel (expressed as a matrix). The eigenvalues are gains of the fundamental modes of the the channel, which themselves are captured by the eigenvectors[19].

Google also uses the eigenvector corresponding to the maximal eigenvalues of the Google matrix to determine the rank of a page for search and according to their ranking, the web-pages are displayed[15].

This thesis is organised as follows:

- Chapter 2: In this chapter we consider bounding the eigenvalues by using the matrix elements. Amongst others we consider the Gershgorin circle theorem and its extension to the ovals of Cassini. The latter two are compared by examples.

- Chapter 3: We concentrate on bounds by matrix norms. In particular we use the Frobenius, infinity and spectral norms. We also consider matrices that are block partitioned.

- Chapter 4: Since eigenvalues are intimately related to the trace of a matrix, in this chapter we consider bounds by traces. Here we expand in some detail a paper by Wolkowicz H and Styan G.P.H[22].

- Chapter 5: In this chapter we consider special tridiagonal matrices which naturally arise by discretization of boundary value problems using finite differences. Here we also expand in some detail the paper by Buchholzer H and Kanzow C[4].

- Chapter 6: Conclusion.

Chapter 2

Bounds by Matrix Elements

Theorem 2.1. [10] *Let A be a complex matrix of order n with conjugate transpose A^* and eigenvalue λ. Define*

$$G = \frac{1}{2}(A + A^*)$$
$$T = \frac{1}{2}(A - A^*)$$

and $a = \max_{i,j} |a_{ij}|, \ g = \max_{i,j} |g_{ij}|, \ t = \max_{i,j} |t_{ij}|, \ \lambda = \alpha + i\beta.$

Then $|\lambda| \leq na, |\alpha| \leq ng$ *and* $|\beta| \leq nt.$

Proof. Let $A\mathbf{x} = \lambda\mathbf{x}$, where $\mathbf{x}$ is a normalized eigenvector such that $\langle \mathbf{x}, \mathbf{x} \rangle = 1$ so that

$$\begin{aligned} \langle A\mathbf{x}, \mathbf{x} \rangle &= \langle \lambda\mathbf{x}, \mathbf{x} \rangle \\ &= \lambda \langle \mathbf{x}, \mathbf{x} \rangle \\ &= \lambda, \end{aligned} \tag{2.1}$$

or equivalently by conjugating (2.1)

$$\langle A^*\mathbf{x}, \mathbf{x} \rangle = \bar{\lambda}.$$

Then

$$\langle A\mathbf{x}, \mathbf{x}\rangle + \langle A^*\mathbf{x}, \mathbf{x}\rangle = \lambda + \bar{\lambda} \tag{2.2}$$

$$= 2\alpha, \tag{2.3}$$

which implies

$$\langle (A + A^*)\mathbf{x}, \mathbf{x}\rangle = 2\alpha.$$

Thus

$$\langle G\mathbf{x}, \mathbf{x}\rangle = \alpha. \tag{2.4}$$

Likewise

$$\langle A\mathbf{x}, \mathbf{x}\rangle - \langle A^*\mathbf{x}, \mathbf{x}\rangle = \langle (A - A^*)\mathbf{x}, \mathbf{x}\rangle \tag{2.5}$$

$$= 2i\beta. \tag{2.6}$$

Thus

$$\langle T\mathbf{x}, \mathbf{x}\rangle = i\beta, \tag{2.7}$$

or equivalently

$$-i\langle T\mathbf{x}, \mathbf{x}\rangle = \beta, \tag{2.8}$$

so

$$\begin{aligned} |\lambda| &= |\langle A\mathbf{x}, \mathbf{x}\rangle| \\ &= \left| \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_j x_i \right| \\ &\leq \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}||x_j||x_i|. \end{aligned} \tag{2.9}$$

From the inequality

$$|x_i|^2 + |x_j|^2 - 2|x_i||x_j| \geq 0, \tag{2.10}$$

we obtain the result

$$|x_i||x_j| \leq \frac{1}{2}(|x_i|^2 + |x_j|^2), \tag{2.11}$$

using (2.11) in (2.9) yields

$$\begin{aligned}
|\lambda| &\leq a\left(\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i||x_j|\right) && (2.12)\\
&\leq \frac{a}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(|x_i|^2 + |x_j|^2\right) && (2.13)\\
&= \frac{a}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i|^2 + \frac{a}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}|x_j|^2 && (2.14)\\
&= \frac{na}{2} + \frac{na}{2} && (2.15)\\
&= na, && (2.16)
\end{aligned}$$

where we have used the fact that $\|\mathbf{x}\| = 1$.

Proceeding in a similar manner, from (2.4) and (2.8) we obtain

$$|\alpha| \leq ng,$$

and

$$|\beta| \leq nt.$$

□

Theorem 2.2. [10] *Let A be a real matrix of order n and T be its skew symmetric part given by*

$$T = \frac{1}{2}(A - A^T)$$

and $\lambda = \alpha + i\beta$ *be an eigenvalue of* A. *Then*

$$|\beta| \le t\sqrt{\frac{n(n-1)}{2}}$$

where $t = \max_{i,j} |t_{ij}|$.

Proof. Since, $A\mathbf{x} = \lambda\mathbf{x}$ for $\mathbf{x} = \mathbf{y} + i\mathbf{z}, i\mathbf{x} = i\mathbf{y} - \mathbf{z}$,

$$A(\mathbf{y} + i\mathbf{z}) = (\alpha + i\beta)(\mathbf{y} + i\mathbf{z}) \tag{2.17}$$

$$= (\alpha\mathbf{y} - \beta\mathbf{z}) + i(\alpha\mathbf{z} + \beta\mathbf{y}). \tag{2.18}$$

Equating imaginary and real part of (2.18) yields

$$A\mathbf{z} = \alpha\mathbf{z} + \beta\mathbf{y} \tag{2.19}$$

and

$$A\mathbf{y} = \alpha\mathbf{y} - \beta\mathbf{z} \tag{2.20}$$

Taking the inner-product in (2.19) and (2.20) with $\mathbf{y}$ and $\mathbf{z}$ respectively yields

$$\langle A\mathbf{z}, \mathbf{y}\rangle = \langle \alpha\mathbf{z}, \mathbf{y}\rangle + \langle \beta\mathbf{y}, \mathbf{y}\rangle \tag{2.21}$$

and

$$-\langle A\mathbf{y}, \mathbf{z}\rangle = -\langle \alpha\mathbf{y}, \mathbf{z}\rangle + \langle \beta\mathbf{z}, \mathbf{z}\rangle. \tag{2.22}$$

By adding equation (2.21) and (2.22) we obtain,

$$\langle A\mathbf{z}, \mathbf{y}\rangle - \langle A\mathbf{y}, \mathbf{z}\rangle = \beta(\langle\ \mathbf{y}, \mathbf{y}\rangle + \langle\ \mathbf{z}, \mathbf{z}\rangle) \tag{2.23}$$

$$= \beta\left(\|\mathbf{y}\|^2 + \|\mathbf{z}\|^2\right) \tag{2.24}$$

Now

$$\langle A\mathbf{z}, \mathbf{y}\rangle - \langle A\mathbf{y}, \mathbf{z}\rangle = \langle A\mathbf{z}, \mathbf{y}\rangle - \langle A^T\mathbf{z}, \mathbf{y}\rangle \tag{2.25}$$

$$= \langle (A - A^T)\mathbf{z}, \mathbf{y}\rangle. \tag{2.26}$$

From equation (2.24) and (2.26) we obtain

$$\langle (A - A^T)\mathbf{z}, \mathbf{y}\rangle = \beta(\|\mathbf{y}\|^2 + \|\mathbf{z}\|^2) \tag{2.27}$$

By definition of T, it follows that

$$\frac{\beta}{2}(\|\mathbf{z}\|^2 + \|\mathbf{y}\|^2) = \langle T\mathbf{z}, \mathbf{y}\rangle \tag{2.28}$$

Since $T = -T^T, t_{ij} = -t_{ji}$ and $t_{ii} = 0$ and also T can be written $T = U - U^T$, where U is a strictly upper triangle matrix. Thus

$$\langle T\mathbf{z}, \mathbf{y}\rangle = \langle (U - U^T)\mathbf{z}, \mathbf{y}\rangle \tag{2.29}$$

$$= \langle U\mathbf{z}, \mathbf{y}\rangle - \langle U^T\mathbf{z}, \mathbf{y}\rangle \tag{2.30}$$

$$= \langle U\mathbf{z}, \mathbf{y}\rangle - \langle U\mathbf{y}, \mathbf{z}\rangle \tag{2.31}$$

$$= \sum_{i=1}^{n} \sum_{j=i+1}^{n} u_{ij} z_j y_i - \sum_{i=1}^{n} \sum_{j=i+1}^{n} u_{ij} y_j z_i \tag{2.32}$$

$$= \sum_{i=1}^{n} \sum_{j=i+1}^{n} u_{ij}(z_j y_i - y_j z_i) \tag{2.33}$$

$$= \sum_{i=1}^{n} \sum_{j=i+1}^{n} t_{ij}(z_j y_i - y_j z_i) \tag{2.34}$$

$$\leq \sum_{i=1}^{n} \sum_{j=i+1}^{n} |t_{ij}||z_j y_i - y_j z_i| \tag{2.35}$$

$$\leq t \sum_{i=1}^{n} \sum_{j=i+1}^{n} |z_j y_i - y_j z_i|. \tag{2.36}$$

Hence (2.28) becomes

$$\frac{\beta}{2}(\|\mathbf{z}\|^2 + \|\mathbf{y}\|^2) \leq t \sum_{i=1}^{n} \sum_{j=i+1}^{n} |z_j y_i - y_j z_i| \tag{2.37}$$

and squaring both sides of (2.37), yields

$$\beta^2(\|\mathbf{z}\|^2 + \|\mathbf{y}\|^2)^2 \leq 4t^2 \left(\sum_{i=1}^{n} \sum_{j=i+1}^{n} |z_j y_i - y_j z_i| \right)^2 \tag{2.38}$$

In order to obtain a bound we prove that for real numbers r_l

$$\left(\sum_{l=1}^{m} r_l \right)^2 \leq m \sum_{l=1}^{m} r_l^2. \tag{2.39}$$

We begin by observing that

$$\left(\sum_{l=1}^{m} r_l \right)^2 = \sum_{p=1}^{m} \sum_{k=1}^{m} r_k r_p$$

and from $(r_k - r_p)^2 \geq 0$ we get,

$$r_k r_p \leq \frac{1}{2}(r_k^2 + r_p^2), \tag{2.40}$$

which implies

$$\begin{aligned} \sum_{p=1}^{m} \sum_{k=1}^{m} r_k r_p &\leq \frac{1}{2} \sum_{p=1}^{m} \sum_{k=1}^{m} (r_k^2 + r_p^2) && (2.41) \\ &= \frac{1}{2} \sum_{p=1}^{m} r_p^2 \sum_{k=1}^{m} (1) + \frac{1}{2} \sum_{k=1}^{m} r_k^2 \sum_{p=1}^{m} (1) && (2.42) \\ &= \frac{m}{2} \sum_{p=1}^{m} r_p^2 + \frac{m}{2} \sum_{k=1}^{m} r_k^2 && (2.43) \\ &= m \sum_{l=1}^{m} r_l^2. && (2.44) \end{aligned}$$

We also observe that the number of off-diagonal elements in U is $\frac{n^2-n}{2}$ or $\frac{n(n-1)}{2}$. From this fact and using (2.39) we get

$$\left(\sum_{i=1}^{n} \sum_{j=i+1}^{n} |z_j y_i - y_j z_i| \right)^2 \leq \frac{n(n-1)}{2} \sum_{i=1}^{n} \sum_{j=i+1}^{n} |z_j y_i - y_j z_i|^2 \tag{2.45}$$

Considering

$$\|\mathbf{y}\|^2 \|\mathbf{z}\|^2 - \frac{1}{4} \left(\|\mathbf{y}\|^2 + \|\mathbf{z}\|^2 \right)^2 = -\frac{1}{4} \left(\|\mathbf{y}\|^2 - \|\mathbf{z}\|^2 \right)^2 \tag{2.46}$$

and Lagrange's identity[21]

$$\|\mathbf{y}\|^2 \|\mathbf{z}\|^2 = \langle \mathbf{y}, \mathbf{z} \rangle^2 + \sum_{i=1}^{n} \sum_{j=i+1}^{n} |z_j y_i - y_j z_i|^2 \tag{2.47}$$

and subtracting (2.46) from (2.47) gives

$$\frac{1}{4}\left(\|\mathbf{y}\|^2 + \|\mathbf{z}\|^2\right)^2 = \langle \mathbf{y}, \mathbf{z} \rangle^2 + \frac{1}{4}\left(\|\mathbf{y}\|^2 - \|\mathbf{z}\|^2\right)^2 + \sum_{i=1}^{n} \sum_{j=i+1}^{n} |z_j y_i - y_j z_i|^2 \tag{2.48}$$

$$\geq \sum_{i=1}^{n} \sum_{j=i+1}^{n} |z_j y_i - y_j z_i|^2. \tag{2.49}$$

Thus

$$\left(\|\mathbf{y}\|^2 + \|\mathbf{z}\|^2\right)^2 \geq 4 \sum_{i=1}^{n} \sum_{j=i+1}^{n} |z_j y_i - y_j z_i|^2. \tag{2.50}$$

Multiply (2.50) by β^2 and use (2.38) and (2.45) to obtain

$$\begin{aligned} 4\beta^2 \sum_{i=1}^{n} \sum_{j=i+1}^{n} |z_j y_i - y_j z_i|^2 &\leq \beta^2 \left(\|\mathbf{y}\|^2 + \|\mathbf{z}\|^2\right)^2 \\ &\leq 4t^2 \left(\sum_{i=1}^{n} \sum_{j=i+1}^{n} |z_j y_i - y_j z_i| \right)^2 \\ &\leq 4t^2 \left[\frac{n(n-1)}{2}\right] \sum_{i=1}^{n} \sum_{j=i+1}^{n} |z_j y_i - y_j z_i|^2 \end{aligned}$$

Thus

$$\beta^2 \leq t^2 \left[\frac{n(n-1)}{2}\right] \tag{2.51}$$

Taking the square root of (2.51) yields

$$|\beta| \leq t\sqrt{\frac{n(n-1)}{2}}.$$

□

Theorem 2.3. [3] *Every eigenvalue of a matrix is contained in at least one of the n disks whose centres are a_{kk} and whose radii are*

$$r_k = \sum_{\substack{m=1 \\ m \neq k}}^{n} |a_{km}| \qquad (k = 1, \cdots, n)$$

Proof. Let B be a matrix of order n. The system of equations $B\mathbf{x} = \mathbf{0}$ has a non-trivial solution if and only if $\det B = 0$. Let x_k be the dominant component of $\mathbf{x} = [x_1, ..., x_n]^T$, that is, $|x_k| \geq |x_m|$ for all m. Then, the k_{th} equation is

$$
\begin{aligned}
b_{kk}x_k &= -\sum_{m\neq k}^{n} b_{km}x_m \\
|b_{kk}||x_k| &= \left|\sum_{m\neq k}^{n} b_{km}x_m\right| \\
&\leq |x_m| \sum_{m\neq k}^{n} |b_{km}| \\
&\leq |x_k| \sum_{m\neq k}^{n} |b_{km}|
\end{aligned}
$$

thus

$$|b_{kk}| \leq \sum_{m\neq k}^{n} |b_{km}|$$

Let $B = A - \lambda I$, where λ is such that $\det(A - \lambda I) = 0$, therefore

$$
\begin{aligned}
|\lambda - a_{kk}| &\leq \sum_{m\neq k}^{n} |a_{km}| \\
&= r_k
\end{aligned}
$$

□

Corollary 2.3.1. *Theorem of Frobenius :*

$$
\begin{aligned}
|\lambda|_{\max} &\leq \max_k \sum_{m=1}^{n} |a_{km}| \\
|\lambda|_{\min} &\geq \min_k \left(|a_{kk}| - \sum_{m\neq k}^{n} |a_{km}| \right)
\end{aligned}
$$

Proof.

$$|\lambda| = |\lambda - a_{kk} + a_{kk}| \tag{2.52}$$

$$\leq |\lambda - a_{kk}| + |a_{kk}| \tag{2.53}$$

$$\leq \sum_{m \neq k}^{n} |a_{km}| + |a_{kk}| \tag{2.54}$$

Hence

$$|\lambda|_{\max} \leq \max_k \sum_{m=1}^{n} |a_{km}| \tag{2.55}$$

$$|\lambda| = |\lambda - a_{kk} - (-a_{kk})| \tag{2.56}$$

$$\geq |a_{kk}| - |\lambda - a_{kk}| \tag{2.57}$$

$$\geq |a_{kk}| - \sum_{m \neq k}^{n} |a_{km}| \tag{2.58}$$

Hence

$$|\lambda|_{\min} \geq \min_k \left(|a_{kk}| - \sum_{m \neq k}^{n} |a_{km}| \right) \tag{2.59}$$

□

Theorem 2.4. [2] *As a further refinement of Theorem 2.3, consider the ovals of Cassini (which restricts the regions containing the eigenvalues). Each eigenvalue of A lies in at least one of the* $\frac{n(n-1)}{2}$ *ovals of Cassini*

$$|\lambda - a_{kk}||\lambda - a_{ll}| \leq \sum_{j \neq k}^{n} |a_{kj}| \sum_{j \neq l}^{n} |a_{lj}|$$

Proof. For $\mathbf{x} = [x_1, x_2, ..., x_n]^T$

$$(\lambda - a_{ii})x_i = \sum_{j \neq i}^{n} a_{ij}x_j$$

Let $|x_k| \geq |x_l| \geq |x_j|$ for $j \neq k, j \neq l$.

Then

$$|\lambda - a_{kk}||x_k| \leq \sum_{j \neq k}^{n} |a_{kj}||x_j| \tag{2.60}$$
$$\leq \left(\sum_{j \neq k}^{n} |a_{kj}| \right) |x_l| \tag{2.61}$$

and

$$|\lambda - a_{ll}||x_l| \leq \sum_{j \neq l}^{n} |a_{lj}||x_j| \tag{2.62}$$
$$\leq \left(\sum_{j \neq l}^{n} |a_{lj}| \right) |x_k| \tag{2.63}$$

Therefore, multiplying (2.61) and (2.63) gives

$$|\lambda - a_{kk}||\lambda - a_{ll}||x_k||x_l| \leq \left(\sum_{j \neq k}^{n} |a_{kj}| \right) \left(\sum_{j \neq l}^{n} |a_{lj}| \right) |x_k||x_l|, \tag{2.64}$$

so that

$$|\lambda - a_{kk}||\lambda - a_{ll}| \leq \left(\sum_{j \neq k}^{n} |a_{kj}| \right) \left(\sum_{j \neq l}^{n} |a_{lj}| \right) \tag{2.65}$$

which proves the theorem and this gives result for row sums. Similarly for column sums it can be shown that

$$|\lambda - a_{kk}||\lambda - a_{ll}| \leq \left(\sum_{i \neq k}^{n} |a_{ik}| \right) \left(\sum_{i \neq l}^{n} |a_{il}| \right) \tag{2.66}$$

□

Another inequality giving the regions in which the eigenvalues are contained is presented in the following theorem.

Theorem 2.5. [18] *For the matrix* $A = (a_{ij})$,

$$|\lambda - a_{ii}| \leq \left(\sum_{j \neq i}^{n} |a_{ij}| \right)^{\alpha} \left(\sum_{k \neq i}^{n} |a_{ki}| \right)^{1-\alpha}$$

for $0 \leq \alpha \leq 1$.

Proof. As it was shown in Theorem 2.3, for the $\det(A - \lambda I)$ to vanish, the following inequalities must be satisfied:

$$|\lambda - a_{ii}| \leq \sum_{j \neq i}^{n} |a_{ij}| \tag{2.67}$$

$$|\lambda - a_{ii}| \leq \sum_{k \neq i}^{n} |a_{ki}| \tag{2.68}$$

Inequality (2.68) arises by considering $x^T A^T = \lambda x^T$.

Thus

$$|\lambda - a_{ii}| = |\lambda - a_{ii}|^{\alpha} |\lambda - a_{ii}|^{1-\alpha} \tag{2.69}$$

$$\leq \left(\sum_{j \neq i}^{n} |a_{ij}| \right)^{\alpha} \left(\sum_{k \neq i} |a_{ki}| \right)^{1-\alpha} \tag{2.70}$$

whenever $0 \leq \alpha \leq 1$. $\square$

Two corollaries result form Theorem 2.5 and are presented below.

Corollary 2.5.1.

$$|\lambda|_{\max} \leq \left[|a_{ii}| + \left(\sum_{j \neq i}^{n} |a_{ij}| \right)^{\alpha} \left(\sum_{k \neq i}^{n} |a_{ki}| \right)^{1-\alpha} \right] \tag{2.71}$$

$$|\lambda|_{\min} \geq \left[|a_{ii}| - \left(\sum_{j \neq i}^{n} |a_{ij}| \right)^{\alpha} \left(\sum_{k \neq i}^{n} |a_{ki}| \right)^{1-\alpha} \right] \tag{2.72}$$

$$|\lambda|_{\max} \leq \left(|a_{ii}| + \sum_{j \neq i}^{n} |a_{ij}| \right)^{\alpha} \left(|a_{ii}| + \sum_{k \neq i}^{n} |a_{ki}| \right)^{1-\alpha} \tag{2.73}$$

for all $0 \leq \alpha \leq 1$.

Proof. The first two inequalities, that is, inequality (2.71) and (2.72) follows from Theorem 2.5.

And for inequality (2.73), we define the summations

$$S_1 = \sum_{j \neq i}^{n} |a_{ij}|, \tag{2.74}$$

$$S_2 = \sum_{k \neq i} |a_{ki}| \tag{2.75}$$

and consider the 2-element sequences $(|a_{ii}|, S_1)$, $(|a_{ii}|, S_2)$ then by Hölder's inequality[17] we have then,

$$|a_{ii}|^{\alpha}|a_{ii}|^{1-\alpha} + S_1^{\alpha} S_2^{1-\alpha} \leq (|a_{ii}| + S_1)^{\alpha} (|a_{ii}| + S_2)^{1-\alpha} \tag{2.76}$$

$$|a_{ii}| + S_1^{\alpha} S_2^{1-\alpha} \leq (|a_{ii}| + S_1)^{\alpha} (|a_{ii}| + S_2)^{1-\alpha} \tag{2.77}$$

$$\tag{2.78}$$

Inequality (2.71) now follows from (2.61).

□

Corollary 2.5.2. *For each α, $0 \leq \alpha \leq 1$, every eigenvalue of A lies in at least one of the $\frac{n(n-1)}{2}$ ovals,*

$$|\lambda - a_{ii}||\lambda - a_{jj}| \leq \left[\left(\sum_{j \neq i}^{n} |a_{ij}|\right)\left(\sum_{k \neq j}^{n} |a_{kj}|\right)\right]^{1-\alpha} \left[\left(\sum_{i \neq j}^{n} |a_{ji}|\right)\left(\sum_{j \neq k}^{n} |a_{jk}|\right)\right]^{\alpha}$$

Proof. This corollary is a direct consequence of (2.70) and for $\alpha = 0$ or $\alpha = 1$, this relation reduces to Theorem 2.4. □

Theorem 2.6. [2]*Each eigenvalue λ satisfies*

$$|\lambda| \leq \frac{1}{2} \max_{k,j=1,2,\ldots,m} \left[|a_{kk}| + |a_{jj}| + \sqrt{(|a_{kk}| - |a_{jj}|)^2 + 4P_k P_j}\right] = M$$

where

$$P_k = \sum_{j \neq k} |a_{kj}|$$

Proof. We assume that $|a_{rr}| \leq |a_{ss}|$.

Case 1: If $|\lambda| \leq |a_{rr}|$, then

$$\begin{aligned} |\lambda| &\leq \tfrac{1}{2}\left(|a_{rr}| + |a_{ss}|\right) + \tfrac{1}{2}(|a_{rr}| - |a_{ss}|) \\ &\leq \tfrac{1}{2}\left[|a_{rr}| + |a_{ss}| + \sqrt{(|a_{rr}| - |a_{ss}|)^2 + 4P_rP_s}\right] \\ &\leq M \end{aligned}$$

This is trivially true, since $|a_{rr}| - |a_{ss}| \leq 0$ and P_r, P_s are positive.

Case 2: If $|\lambda| > |a_{rr}| \geq |a_{ss}|$, then from the triangle inequality it follows that

$$0 < |\lambda| - |a_{rr}| \leq |\lambda - a_{rr}| \tag{2.79}$$

and

$$0 < |\lambda| - |a_{ss}| \leq |\lambda - a_{ss}| \tag{2.80}$$

Producting (2.79) and (2.80) and using Theorem 2.4 gives

$$\begin{aligned} (|\lambda| - |a_{rr}|)(|\lambda| - |a_{ss}|) &\leq |\lambda - a_{rr}||\lambda - a_{ss}| \\ &\leq P_rP_s \end{aligned} \tag{2.81}$$

$$|\lambda|^2 - (|a_{rr}| + |a_{ss}|)|\lambda| + |a_{rr}||a_{ss}| - P_rP_s \leq 0 \tag{2.82}$$

Solving the quadratic equation resulting from (2.82) we get the zeroes

$$|\lambda_1| = \tfrac{1}{2}\left[|a_{rr}| + |a_{ss}| + \sqrt{(|a_{rr}| - |a_{ss}|)^2 + 4P_rP_s}\right] \tag{2.83}$$

and

$$|\lambda_2| = \tfrac{1}{2}\left[|a_{rr}| + |a_{ss}| - \sqrt{(|a_{rr}| - |a_{ss}|)^2 + 4P_rP_s}\right] \tag{2.84}$$

From Figure 2.1 representing the quadratic from (2.82) we observe that

$$
\begin{aligned}
|\lambda| &\leq |\lambda_1| \\
&\leq M
\end{aligned}
$$

□

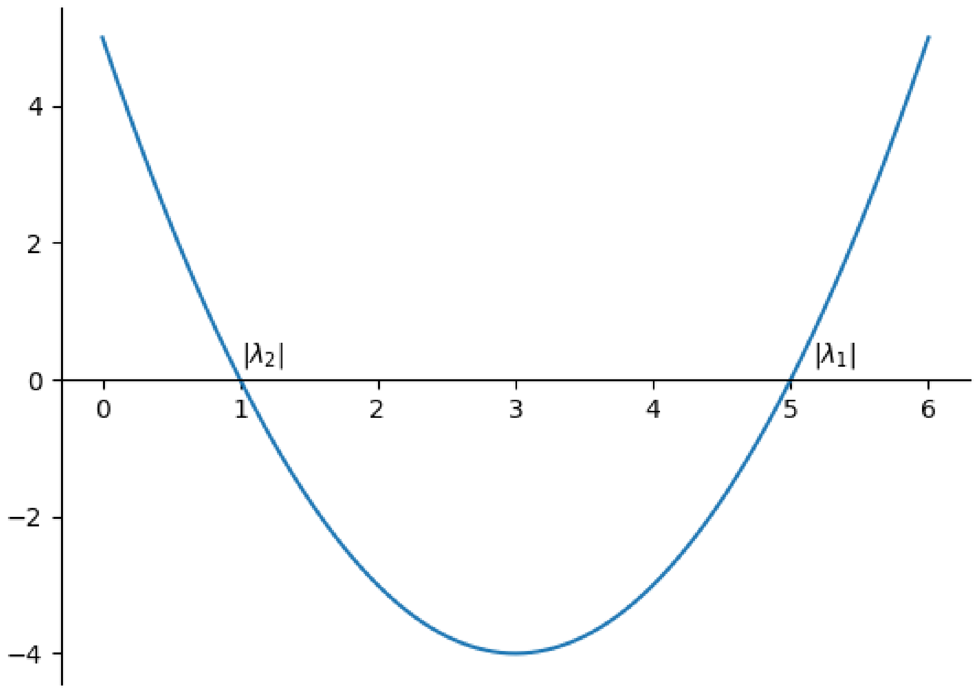


Figure 2.1: Quadratic from (2.82)

Theorem 2.7. [2] *If*

$$|a_{kk}a_{jj}| > P_k P_j$$

then

$$|\lambda| \geq \tfrac{1}{2} \min_{k,j=1,2,\ldots,n} \left[|a_{kk}| + |a_{jj}| - \sqrt{(|a_{kk}| - |a_{jj}|)^2 + 4P_kP_j} \right]$$

$$= m > 0$$

Proof. From theorem 2.6 it follows that

$$|\lambda| \geq |\lambda_2| \tag{2.85}$$

$$= \tfrac{1}{2} \left[|a_{rr}| + |a_{ss}| - \sqrt{(|a_{rr}| - |a_{ss}|)^2 + 4P_rP_s} \right] \tag{2.86}$$

$$\geq m \tag{2.87}$$

Assume that m is attained where $k = f$ and $j = g$

$$m = \tfrac{1}{2} \left[|a_{ff}| + |a_{gg}| - \sqrt{(|a_{ff}| - |a_{gg}|)^2 + 4P_fP_g} \right] \tag{2.88}$$

$$= \tfrac{1}{2} \left[|a_{ff}| + |a_{gg}| - \sqrt{|a_{ff}|^2 - 2|a_{ff}||a_{gg}| + |a_{gg}|^2 + 4P_fP_g} \right] \tag{2.89}$$

$$> \tfrac{1}{2} \left[|a_{ff}| + |a_{gg}| - \sqrt{|a_{ff}|^2 - 2|a_{ff}||a_{gg}| + |a_{gg}|^2 + 4|a_{ff}||a_{gg}|} \right] \tag{2.90}$$

$$= \tfrac{1}{2} \left[|a_{ff}| + |a_{gg}| - \sqrt{(|a_{ff}| + |a_{gg}|)^2} \right] \tag{2.91}$$

$$= 0 \tag{2.92}$$

□

2.1 Examples

Example 1. Here we tested result of Theorem 2.1 by using;

$$A = \begin{bmatrix} 1+i & 2 & 3-2i & 5 \\ 6+3i & 4 & 1-i & 8 \\ 2-i & 4 & 3+i & 5 \\ 1+i & 1-i & 2-i & 2+i \end{bmatrix} \tag{2.93}$$

then

$$G = \begin{bmatrix} 1 & 4.0-1.5i & 2.5-0.5i & 3.0-0.5i \\ 4.0+1.5i & 4.0+0.0i & 2.5-0.5i & 4.5+0.5i \\ 2.5+0.5i & 2.5+0.5i & 3 & 3.5+0.5i \\ 3.0+0.5i & 4.5-0.5i & 3.5-0.5i & 2 \end{bmatrix} \tag{2.94}$$

and

$$T = \begin{bmatrix} 1.0i & -2.0+1.5i & 0.5-1.5i & 2.0+0.5i \\ 2.0+1.5i & 0 & -1.5-0.5i & 3.5-0.5i \\ -0.5-1.5i & 1.5-0.5i & 1.0i & 1.5-0.5i \\ -2.0+0.5i & -3.5-0.5i & -1.5-0.5i & 1.0i \end{bmatrix} \tag{2.95}$$

Hence $a = 8,\ g = 4.5277,\ t = 3.5355$. Table 2.1 shows calculated values and we observe from Theorem 2.1 that $|\lambda| \leq 32,\ |\alpha| \leq 18.1108$ and $|\beta| \leq 14.1421$, which are rather large upper bounds.

Example 2. Now we test the result of Theorem 2.2 by using;

$$B = \begin{bmatrix} 1 & 2 & 9 & 4 \\ 3 & 2 & 1 & 4 \\ 1 & 1 & 2 & 2 \\ 3 & 7 & 1 & 1 \end{bmatrix} \tag{2.96}$$

then

$$T = \begin{bmatrix} 0 & -0.5 & 4 & 0.5 \\ 0.5 & 0 & 0 & -1.5 \\ -4 & 0 & 0 & 0.5 \\ -0.5 & 1.5 & -0.5 & 0 \end{bmatrix} \tag{2.97}$$

Hence $t = 4$, Table 2.2 shows calculated values and we observe from Theorem 2.2 that $|\beta| \leq 9.7980$, which again is a large upper bound.

Table 2.1

λ	$\|\lambda\|$	$\|\alpha\|$	$\|\beta\|$
11.5857-0.4246i	11.5951	11.5857	0.4246
-0.5130+4.3984i	4.4282	0.5130	4.3984
-0.3706-1.5266i	1.5709	0.3706	1.5266
-0.7037+0.5528i	0.8949	0.7037	0.5528

Table 2.2

λ	$\|\beta\|$
10.6157	0
-0.4492+0.6760i	0.6760
-0.4492-0.6760i	0.6760
-3.7172	0

Example 3. We consider matrix C

$$C = \begin{bmatrix} 8 & -1 & -5 \\ -4 & 4 & -2 \\ 18 & -5 & -7 \end{bmatrix} \tag{2.98}$$

with eigenvalues $2 \pm 4i$ and 1, which are illustrated by dots in Figure 2.2.

Using matrix C we test Theorem 2.3 and we observed from Figure 2.2a that the eigenvalues lie in the union of Gershgorin disks, using the columns of C. Figure 2.2b illustrate the disks using the rows of C.

Example 4. Again we use matrix C to test the ovals of Cassini corresponding to (2.65) and (2.66), which are depicted in Figures 2.3a and 2.3b.

Plotting ovals of Cassini

Let the equation of the oval be given by

$$|z-a||z-b| = c^2 \tag{2.99}$$

Let $z = x + iy$ and using the transformation

$$z = \hat{z} + \frac{a+b}{2} \tag{2.100}$$

Equation (2.99) becomes

$$|\hat{z}+d||\hat{z}-d| = c^2 \qquad \text{where} \quad d = \tfrac{b-a}{2} \tag{2.101}$$

Squaring (2.101) gives

$$(\hat{z}^2 - d^2)(\bar{\hat{z}}^2 - d^2) = c^4 \tag{2.102}$$

which simplifies to

$$|\hat{z}|^4 - d^2 2Re(\hat{z}^2) + d^4 = c^4 \tag{2.103}$$

Now using $\hat{z} = re^{i\theta}$ we obtain the quartic in r

$$r^4 - 2r^2 d^2 \cos(2\theta) + (d^4 - c^4) = 0 \tag{2.104}$$

Using the quadratic formula we obtain

$$r^2 = \frac{2d^2\cos(2\theta) \pm \sqrt{4d^4\cos^2(2\theta) - 4(d^4 - c^4)}}{2} \tag{2.105}$$

from what we obtain

$$r = d\sqrt{\cos(2\theta) + \sqrt{(\tfrac{c}{d})^4 - \sin^2(2\theta)}} \tag{2.106}$$

from (2.100) we finally get $x = r\cos\theta + \frac{a+b}{2}$ and $y = r\sin\theta$ with $\theta \in [0, 2\pi]$.

Comparing Figure 2.2 and Figure 2.3 we observe that the ovals of Cassini can yield superior bounds.

Example 5. For matrix C, Theorem 2.5 is tested for both $\alpha = \frac{1}{2}$ and $\alpha = \frac{1}{3}$, these are shown in Figure 2.4

Example 6. We consider matrix D

$$D = \begin{bmatrix} 10 & 1 & 2 \\ 3 & 6 & 1 \\ 2 & 5 & 9 \end{bmatrix} \tag{2.107}$$

with eigenvalues 13.0602 and $5.970 \pm 0.651i$.

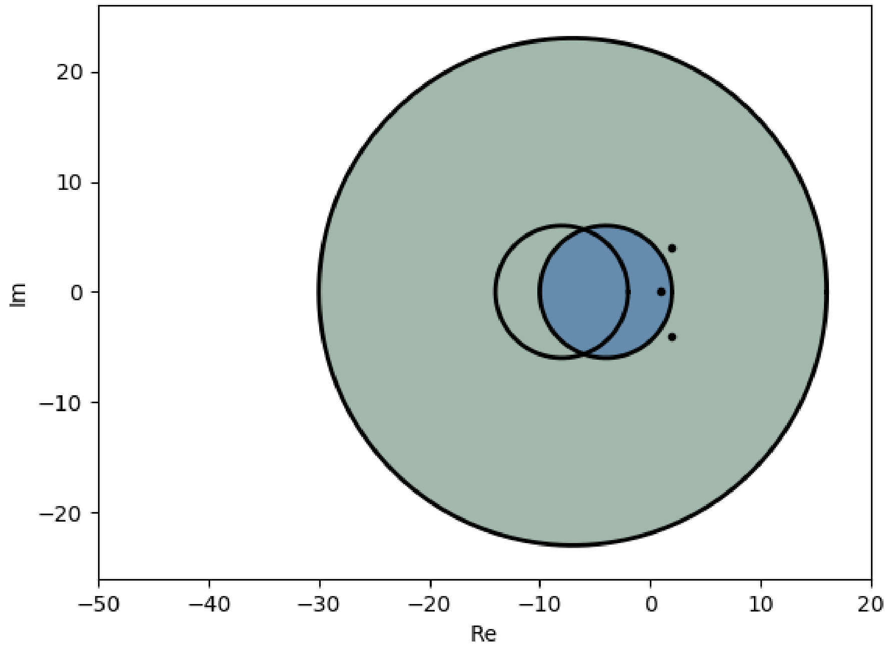
Theorem 2.6 and Theorem 2.7 are tested using matrix D. Let

$$M_{kj} = \frac{1}{2}\left[|a_{kk}| + |a_{jj}| + \sqrt{(|a_{kk}| - |a_{jj}|)^2 + 4P_kP_j}\right]$$

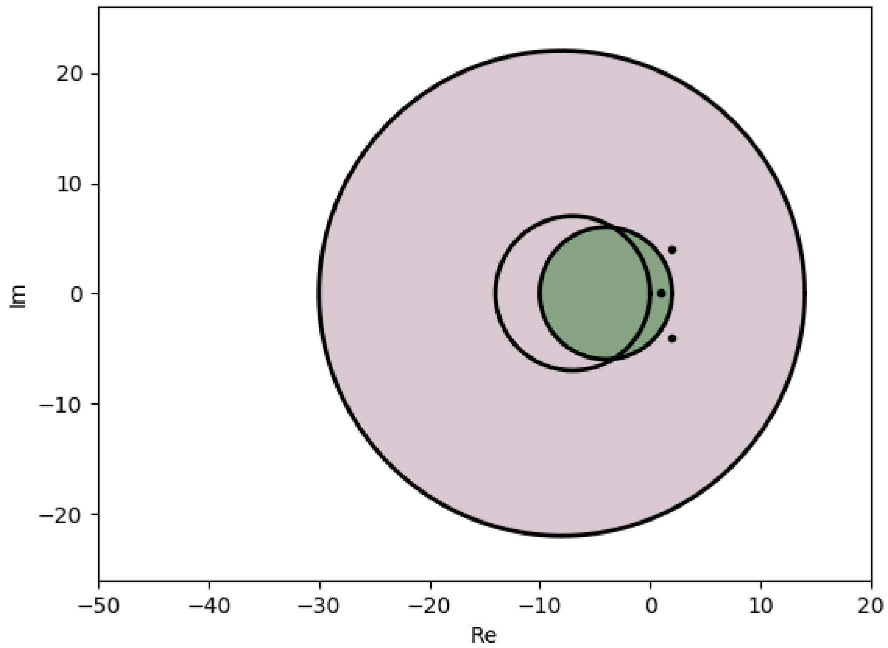
and

$$m_{kj} = \frac{1}{2}\left[|a_{kk}| + |a_{jj}| - \sqrt{(|a_{kk}| - |a_{jj}|)^2 + 4P_kP_j}\right]$$

then these quantities are presented in Table 2.3.

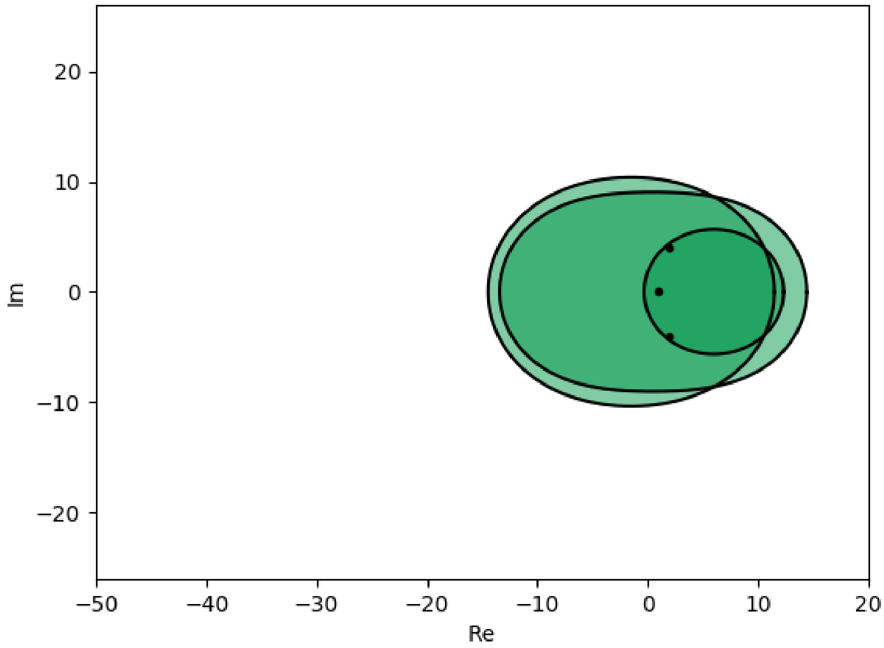


(a) Using columns

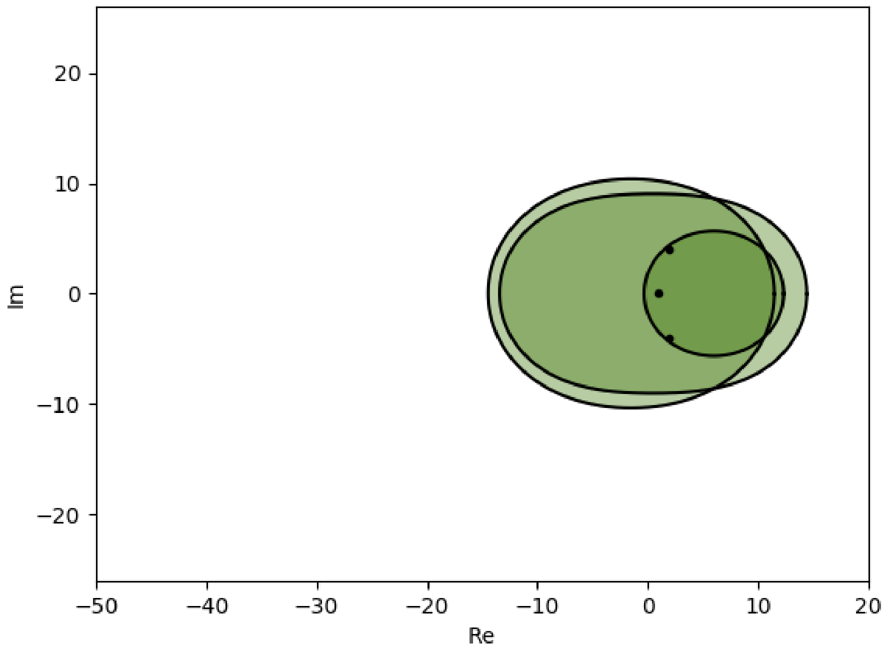


(b) Using rows

Figure 2.2: Gershgorin disks

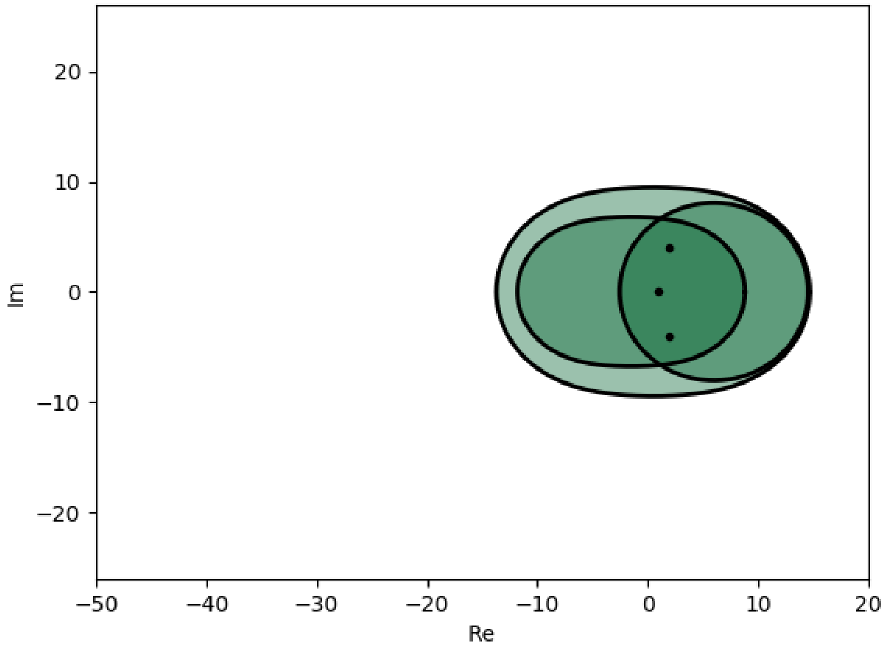


(a) Using columns

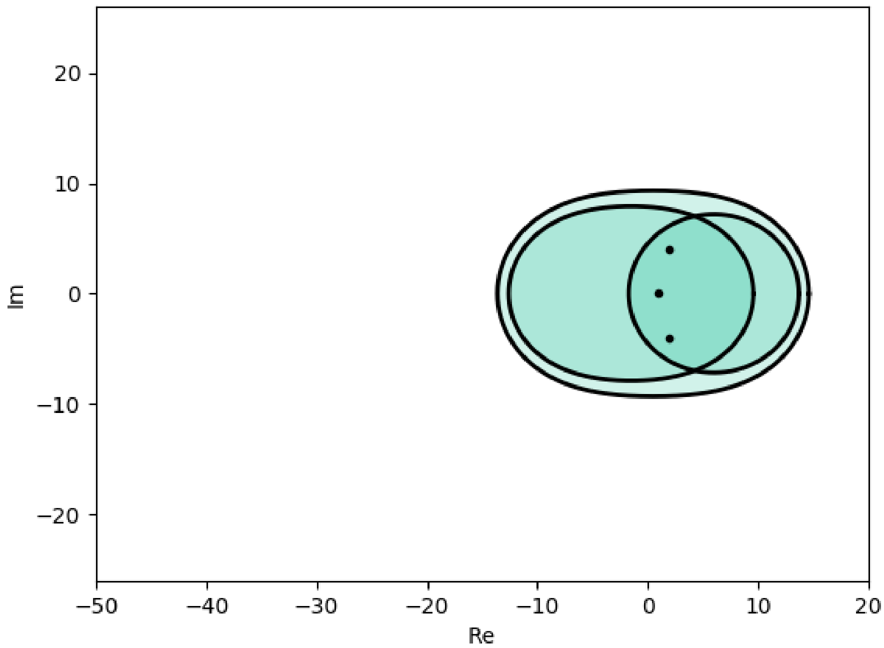


(b) Using rows

Figure 2.3: Ovals of Cassini



(a) $\alpha = \frac{1}{2}$



(b) $\alpha = \frac{1}{3}$

Figure 2.4: Ovals of Cassini

Table 2.3

(k, j)	M_{kj}	m_{kj}
(1,2)	12	4
(1,3)	13.2749	5.7251
(2,3)	13	2

It is observed that $2 = m \leq |\lambda| \leq M = 13.2749$. And we further observe that the upper bound is tight and lower bound is reasonable.

Chapter 3

Bound by Matrix Norms

Definition 3.0.1. A matrix norm $\|\cdot\|$ is a function from $\mathbb{C}^{m\times n}$, the vector space of all matrices of order $m \times n$ to $\mathbb{R}$ which satisfies the following properties for matrices A and B.

1. $\|A\| \geq 0$ whenever $A \neq 0$ and $\|A\| = 0$ iff $A = 0$
2. $\|\alpha A\| = |\alpha|\|A\|$ where α is a scalar
3. $\|A + B\| \leq \|A\| + \|B\|$
4. $\|AB\| \leq \|A\|\|B\|$

Definition 3.0.2. If $\|\cdot\|$ is a vector norm, its induced matrix norm is defined by

$$\|A\| = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}$$

Not all matrix norms are induced by vector norms, in particular the function defined by

$$\|A\|_{\mathbf{F}} = \left(\sum_{i,j} |a_{ij}|^2\right)^{\frac{1}{2}}$$

is indeed a matrix norm called the Frobenius norm. This is now verified

Property 1. The first property is obvious

Property 2.

$$\|\alpha A\|_{\mathbf{F}}^2 = \sum_{i,j} |\alpha a_{ij}|^2 \tag{3.1}$$

$$= |\alpha|^2 \sum_{i,j} |a_{ij}|^2 \tag{3.2}$$

$$= |\alpha|^2 \|A\|_{\mathbf{F}}^2 \tag{3.3}$$

Hence, by square rooting we get $\|\alpha A\|_{\mathbf{F}} = |\alpha| \|A\|_{\mathbf{F}}$.

Property 3. Firstly we observe that

$$|a_{ij} + b_{ij}|^2 = (a_{ij} + b_{ij})(\bar{a}_{ij} + \bar{b}_{ij}) \tag{3.4}$$

$$= |a_{ij}|^2 + |b_{ij}|^2 + a_{ij}\bar{b}_{ij} + \bar{a}_{ij} b_{ij} \tag{3.5}$$

$$= |a_{ij}|^2 + |b_{ij}|^2 + 2\left[Re(a_{ij}\bar{b}_{ij})\right] \tag{3.6}$$

$$\leq |a_{ij}|^2 + |b_{ij}|^2 + 2|a_{ij}\bar{b}_{ij}| \tag{3.7}$$

$$= |a_{ij}|^2 + |b_{ij}|^2 + 2|a_{ij} b_{ij}| \tag{3.8}$$

Thus

$$\|A + B\|_{\mathbf{F}}^2 = \sum_{i,j} |a_{ij} + b_{ij}|^2 \tag{3.9}$$

$$\leq \sum_{i,j} |a_{ij}|^2 + \sum_{i,j} |b_{ij}|^2 + 2\sum_{i,j} |a_{ij} b_{ij}| \tag{3.10}$$

and by using the Cauchy-Schwarz inequality,

$$\|A+B\|_{\mathbf{F}}^2 \leq \|A\|_{\mathbf{F}}^2 + \|B\|_{\mathbf{F}}^2 + 2\left(\sum_{i,j}|a_{ij}|^2\right)^{\frac{1}{2}}\left(\sum_{i,j}|b_{ij}|^2\right)^{\frac{1}{2}} \tag{3.11}$$

$$= \|A\|_{\mathbf{F}}^2 + \|B\|_{\mathbf{F}}^2 + 2\|A\|_{\mathbf{F}}\|B\|_{\mathbf{F}} \tag{3.12}$$

$$= (\|A\|_{\mathbf{F}} + \|B\|_{\mathbf{F}})^2 \tag{3.13}$$

Hence, by square rooting we obtain $\|A+B\|_{\mathbf{F}} \leq \|A\|_{\mathbf{F}} + \|B\|_{\mathbf{F}}$.

Property 4.

$$\|AB\|_{\mathbf{F}}^2 = \sum_{i,j}\sum_{k}|a_{ik}b_{kj}|^2 \tag{3.14}$$

$$\leq \sum_{i,j}\left(\sum_{k}|a_{ik}b_{kj}|\right)^2 \tag{3.15}$$

$$\leq \left(\sum_{i,k}|a_{ik}|^2\right)\left(\sum_{k,j}|b_{kj}|^2\right) \tag{3.16}$$

$$= \|A\|_{\mathbf{F}}^2\|B\|_{\mathbf{F}}^2 \tag{3.17}$$

Statement (3.16) is obtained from (3.15) by using Cauchy-Schwarz inequality. Hence property 4 follows.

Theorem 3.1. [14] *For an arbitrary matrix A, the largest possible eigenvalue modulus is $|\lambda_1| \leq \|A\|$ for any matrix norm of A.*

Proof. Let λ be an eigenvalue of A. Then there is vector $\mathbf{x} \neq \mathbf{0}$ such that $A\mathbf{x} = \lambda\mathbf{x}$. Define the $n \times n$ matrix

$$A_{\mathbf{x}} = [\mathbf{x}, \mathbf{0}, \mathbf{0}, ..., \mathbf{0}] \tag{3.18}$$

then, clearly,

$$AA_{\mathbf{x}} = \lambda A_{\mathbf{x}} \tag{3.19}$$

Using properties (2) and (4) of the matrix norm we deduce that

$$|\lambda|\|A_{\mathbf{x}}\| \leq \|A\|\|A_{\mathbf{x}}\| \tag{3.20}$$

and since $A_{\mathbf{x}} \neq \mathbf{0}$, the first property implies that $\|A_{\mathbf{x}}\| \neq 0$ and hence

$$|\lambda| \leq |\lambda_1| \leq \|A\|. \tag{3.21}$$

$\square$

Theorem 3.2. [14] *If A is $n{\times}n$ matrix, let λ be spectral radius of A^*A, then* $\|A\| = \sqrt{\lambda_1}$

Proof. The matrix A^*A is Hermitian and positive definite since

$$(A^*A)^* = A^*A \tag{3.22}$$

and

$$\mathbf{x}^*(A^*A)\mathbf{x} = (A\mathbf{x})^*A\mathbf{x} = \langle A\mathbf{x}, A\mathbf{x}\rangle \geq 0 \tag{3.23}$$

Let $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$ be a set of orthonormal right-hand eigenvectors of A^*A with associated eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Let $||\cdot||$ denote the Euclidean vector norm and for any $\mathbf{x}$ with $\|\mathbf{x}\| = 1$ we write $\mathbf{x} = \sum\limits_{j=1}^{n} p_j\mathbf{x}_j$. Then

$$A^*A\mathbf{x} = \sum_{j=1}^{n} p_j\lambda_j\mathbf{x}_j \tag{3.24}$$

and

$$\|A\mathbf{x}\|^2 = (A\mathbf{x})^*A\mathbf{x} \tag{3.25}$$

$$= \mathbf{x}^*\left(A^*A\mathbf{x}\right) \tag{3.26}$$

$$= \left(\sum_j p_j\mathbf{x}_j\right)^* \left(\sum_k p_k\lambda_k\mathbf{x}_k\right) \tag{3.27}$$

$$= \sum_j |p_j|^2\lambda_j \tag{3.28}$$

using the fact that $\mathbf{x}_j^* \mathbf{x}_k = \delta_{jk}$.

Thus,

$$\|A\mathbf{x}\| = \left(\sum_j |p_j|^2 \lambda_j \right)^{\frac{1}{2}} \tag{3.29}$$

It follows from (3.28) that

$$\|A\mathbf{x}\|^2 \leq \lambda_1 \sum_j |p_j|^2 \tag{3.30}$$

$$= \lambda_1, \tag{3.31}$$

since $\sum_j |p_j|^2 = 1$. Now

$$\|A\mathbf{x}_1\|^2 = \langle A\mathbf{x}_1, A\mathbf{x}_1 \rangle \tag{3.32}$$

$$= \langle A^* A\mathbf{x}_1, \mathbf{x}_1 \rangle \tag{3.33}$$

$$= \lambda_1 \langle \mathbf{x}_1, \mathbf{x}_1 \rangle \tag{3.34}$$

$$= \lambda_1 \tag{3.35}$$

Inequality (3.31) and equation (3.35) implies that $\|A\| = \sqrt{\lambda_1}$. □

Theorem 3.3. [5] *If A is an $n \times n$ matrix, then*

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|.$$

Proof. Let $\mathbf{x}$ be an n-dimensional vector with $1 = \|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$. Since $A\mathbf{x}$ is also an n-dimensional vector,

$$\|A\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |(A\mathbf{x})_i| \tag{3.36}$$

$$= \max_{1 \leq i \leq n} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \tag{3.37}$$

$$\leq \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}||x_j| \tag{3.38}$$

but $|x_j| \leq \|\mathbf{x}\|_\infty = 1$, hence

$$\|A\mathbf{x}\|_\infty \leq \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}| \|\mathbf{x}\|_\infty. \tag{3.39}$$

Suppose that $\|A\|_\infty = \sum_j |a_{mj}|$ and consider the vector $\tilde{\mathbf{x}}$ with elements defined by

$$\tilde{\mathbf{x}}_j = \begin{cases} \dfrac{|a_{mj}|}{a_{mj}} & \text{if } a_{mj} \neq 0, \\ \\ 0 & \text{if } a_{mj} = 0, \end{cases} \tag{3.40}$$

for $j = (1, 2, \cdots, n)$, $\|\tilde{\mathbf{x}}\|_\infty = 1$, then

$$\|A\tilde{\mathbf{x}}\|_\infty = \max_i \left\| \sum_j a_{ij}\tilde{x}_j \right\| \tag{3.41}$$

$$\geq \left\| \sum_j a_{mj}\tilde{x}_j \right\| \tag{3.42}$$

$$= \sum_j |a_{mj}| \tag{3.43}$$

$$= \sum_j |a_{mj}| \|\tilde{\mathbf{x}}\|_\infty \tag{3.44}$$

$$\frac{\|A\tilde{\mathbf{x}}\|_\infty}{\|\tilde{\mathbf{x}}\|_\infty} \geq \sum_j |a_{mj}| = \|A\|_\infty \tag{3.45}$$

From (3.39)

$$\frac{\|A\tilde{\mathbf{x}}\|_\infty}{\|\tilde{\mathbf{x}}\|_\infty} \leq \|A\|_\infty \tag{3.46}$$

Hence the result follows from (3.45) and (3.46). □

Let matrix A be partitioned such that each diagonal submatrix A_{jj} is square.

Theorem 3.4. [7] *For every such partitioning of the matrix A, each eigenvalue λ of A satisfies*

$$\left\|(A_{jj} - \lambda I_j)^{-1}\right\|^{-1} \leq \sum_{k\neq j}^{N} \|A_{jk}\|$$

whenever the $(A_{jj} - \lambda I_j)^{-1}$ exists.

Proof. Assume that $A - \lambda I$ is singular. Then there exist a nonzero partitioned vector $\mathbf{x} = [\mathbf{x}_1, \cdots, \mathbf{x}_N]^T$such that

$$(A - \lambda I)\mathbf{x} = 0 \tag{3.47}$$

Consider $A - \lambda I$ in its partitioned form, this relation implies

$$\sum_{j \neq i}^{N} A_{ij}\mathbf{x}_j = -(A_{ii} - \lambda I_i)\mathbf{x}_i \tag{3.48}$$

Let $\mathbf{x}_r$ be the largest component of $\mathbf{x}$, in the sense that

$$\|\mathbf{x}_r\| \geq \|\mathbf{x}_j\|, \tag{3.49}$$

then from (3.48) using the r_{th} component

$$\left\| \sum_{j \neq r}^{N} A_{rj}\mathbf{x}_j \right\| = \|(A_{rr} - \lambda I_r)\mathbf{x}_r\|. \tag{3.50}$$

From (3.50) it follows that

$$\|(A_{rr} - \lambda I_r)\mathbf{x}_r\| \leq \sum_{j \neq r}^{N} \|A_{rj}\|\|\mathbf{x}_j\| \leq \sum_{j \neq r}^{N} \|A_{rj}\|\|\mathbf{x}_r\| \tag{3.51}$$

Thus,

$$\frac{\|(A_{rr} - \lambda I_r)\mathbf{x}_r\|}{\|\mathbf{x}_r\|} \leq \sum_{j \neq r}^{N} \|A_{rj}\| \tag{3.52}$$

Now we let

$$\mathbf{z}_{rr} = (A_{rr} - \lambda I_r)\mathbf{x}_r, \tag{3.53}$$

hence

$$\mathbf{x}_r = (A_{rr} - \lambda I_r)^{-1}\mathbf{z}_{rr} \tag{3.54}$$

Substituting (3.53) and (3.54) into (3.52) we get

$$\sum_{j\neq r}^{N} \|A_{rj}\| \geq \frac{\|\mathbf{z}_{rr}\|}{\|(A_{rr}-\lambda I_r)^{-1}\mathbf{z}_{rr}\|} \tag{3.55}$$

Since

$$\frac{\|(A_{rr}-\lambda I_r)^{-1}\mathbf{z}_{rr}\|}{\|\mathbf{z}_{rr}\|} \leq \|(A_{rr}-\lambda I_r)^{-1}\| \tag{3.56}$$

inverting (3.56) we get

$$\frac{\|\mathbf{z}_{rr}\|}{\|(A_{rr}-\lambda I_r)^{-1}\mathbf{z}_{rr}\|} \geq \frac{1}{\|(A_{rr}-\lambda I_r)^{-1}\|} \tag{3.57}$$

It now follows from (3.55) that

$$\sum_{j\neq r}^{N} \|A_{rj}\| \geq \frac{1}{\|(A_{rr}-\lambda I_r)^{-1}\|}.$$

□

If, in theorem 2.4, $|\lambda - a_{ii}|$ is replaced by $\|(A_{ii}-\lambda I_i)^{-1}\|^{-1}$, for $i = k, j$ and $\sum_{j\neq k}^{n} |a_{kj}|$, $\sum_{j\neq l}^{n} |a_{lj}|$ is replaced by $\sum_{j\neq k}^{N} \|A_{kj}\|$, $\sum_{j\neq l}^{N} \|A_{lj}\|$ respectively. Then, the following corollary can be proved.

Corollary 3.4.1. *All eigenvalues of A lie in the union of the $\frac{N(N-1)}{2}$ point sets defined by*

$$\|(A_{kk}-\lambda I_k)^{-1}\|^{-1}\|(A_{ll}-\lambda I_l)^{-1}\|^{-1} \leq \left(\sum_{j\neq k}^{N} \|A_{kj}\|\right)\left(\sum_{j\neq l}^{N} \|A_{lj}\|\right)$$

where $1 \leq k, l \leq N, k \neq l$.

In a similar manner, if these substitutions are made in theorem 2.5, then the following corollary can be proved.

Corollary 3.4.2. *For any α with $0 \leq \alpha \leq 1$, each eigenvalue of A satisfies*

$$\|(A_{ii}-\lambda I_i)^{-1}\|^{-1} \leq \left(\sum_{j\neq i}^{N} \|A_{ij}\|\right)^{\alpha}\left(\sum_{k\neq i}^{N} \|A_{ki}\|\right)^{1-\alpha}$$

for at least one $i, 1 \leq i \leq N$.

3.1 Examples

Consider the partitioned matrix

$$A = \left[\begin{array}{cc|cc} 8 & -4 & -2 & 0 \\ -4 & 8 & 0 & -2 \\ \hline -2 & 0 & 8 & -4 \\ 0 & -2 & -4 & 8 \end{array}\right] = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right]$$

with eigenvalues $\lambda = 2, 6, 10, 14$ and where the vector norm is taken as

$$\|\mathbf{x}\| = \left(\sum_{i=1}^{2} |\mathbf{x}_i|^2 \right)^{1/2} \tag{3.58}$$

for $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2]$. Hence the corresponding matrix norm is the spectral norm. Clearly

$$\|A_{12}\| = \|A_{21}\| = 2 \tag{3.59}$$

Consider

$$(A_{11} - \lambda I) = \begin{bmatrix} 8-\lambda & -4 \\ -4 & 8-\lambda \end{bmatrix}, \tag{3.60}$$

then

$$(A_{11} - \lambda I)^{-1} = \frac{1}{(8-\lambda)^2 - 16} \begin{bmatrix} 8-\lambda & 4 \\ 4 & 8-\lambda \end{bmatrix}$$

Let β denote the eigenvalue of $(A_{11} - \lambda I)^{-1}$, then it is easy to show that

$$\beta \in \left\{ \frac{4-\lambda}{(8-\lambda)^2 - 16}, \frac{12-\lambda}{(8-\lambda)^2 - 16} \right\}, \tag{3.61}$$

(1). If

$$\|(A_{11} - \lambda I)^{-1}\| = \left| \frac{4-\lambda}{(8-\lambda)^2 - 16} \right| \tag{3.62}$$

then from Theorem 3.4

$$\left| \frac{(8-\lambda)^2 - 16}{4-\lambda} \right| \leq 2 \tag{3.63}$$

which simplifies to $|\lambda - 12| \leq 2$.

(2). If

$$\|(A_{11} - \lambda I)^{-1}\| = \left| \frac{12-\lambda}{(8-\lambda)^2 - 16} \right| \tag{3.64}$$

then from Theorem 3.4

$$\left| \frac{(8-\lambda)^2 - 16}{12-\lambda} \right| \leq 2 \tag{3.65}$$

which simplifies to $|\lambda - 4| \leq 2$.

These regions are shown as shaded in Figure 3.1. The largest circle represent the Gershgorin circle.

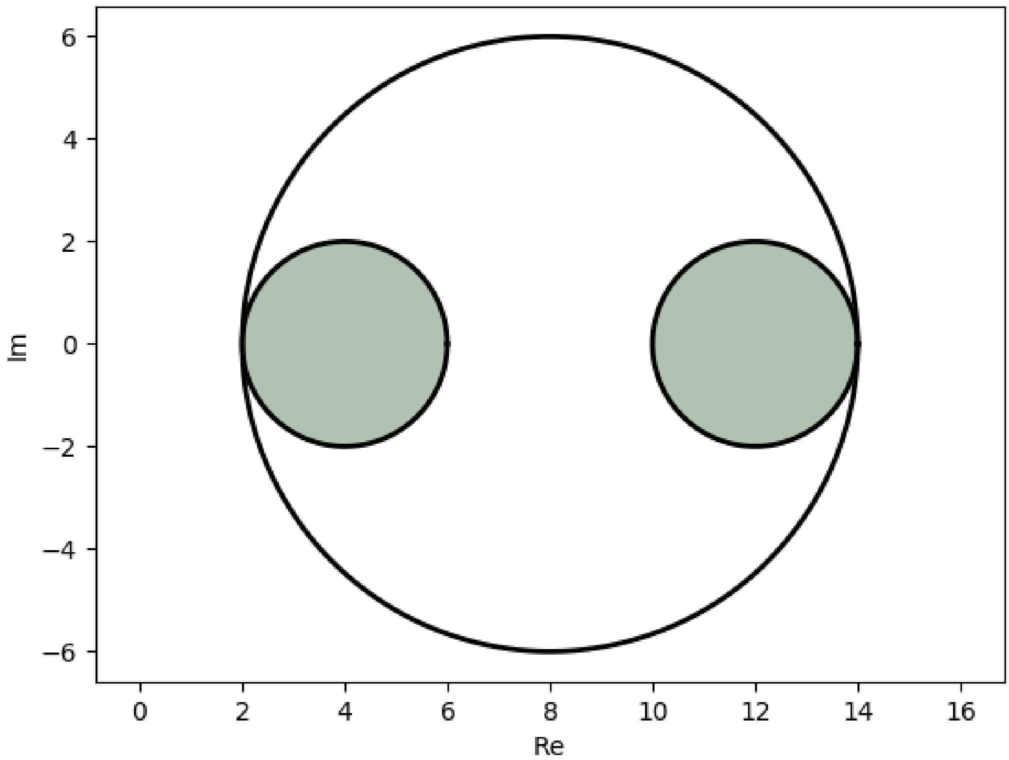


Figure 3.1: Disks from Example 3.1 inside Gershgorin disk

Chapter 4

Bound by Traces

Theorem 4.1. [14] *The eigenvalues of a Hermitian matrix A are real.*

Proof. Let λ be an eigenvalue of A with corresponding eigenvector $\mathbf{x}$ then

$$\begin{aligned}\langle A\mathbf{x}, \mathbf{x}\rangle &= \langle \lambda\ \mathbf{x}, \mathbf{x}\rangle \\ &= \lambda\langle \mathbf{x}, \mathbf{x}\rangle \qquad (4.1)\\ \langle \overline{A\mathbf{x}, \mathbf{x}}\rangle &= \bar{\lambda}\langle \mathbf{x}, \mathbf{x}\rangle \\ \langle \mathbf{x}, A\mathbf{x}\rangle &= \bar{\lambda}\langle \mathbf{x}, \mathbf{x}\rangle \\ \langle A^*\mathbf{x}, \mathbf{x}\rangle &= \bar{\lambda}\langle \mathbf{x}, \mathbf{x}\rangle \qquad (4.2)\end{aligned}$$

Since $A^* = A$ we have,

$$\langle \lambda - \bar{\lambda}\rangle\langle \mathbf{x}, \mathbf{x}\rangle = 0 \qquad (4.3)$$

which implies $\lambda = \bar{\lambda}$. □

Theorem 4.2. [14] *For any matrix A, the trace of A^*A is the same as the square of its Frobenius norm, that is*

$$tr(A^*A) = \|A\|_{\mathbf{F}}^2.$$

Proof.

$$(A^*A)_{ij} = \sum_k a^*_{ik} a_{kj} \tag{4.4}$$

$$= \sum_k \bar{a}_{ki} a_{kj} \tag{4.5}$$

$$(A^*A)_{ii} = \sum_k \bar{a}_{ki} a_{ki} \tag{4.6}$$

$$= \sum_k |a_{ki}|^2. \tag{4.7}$$

From (4.7) it follows that,

$$tr(A^*A) = \sum_i (A^*A)_{ii} \tag{4.8}$$

$$= \sum_i \sum_k |a_{ki}|^2 \tag{4.9}$$

$$= \|A\|^2_{\mathbf{F}} \tag{4.10}$$

□

Theorem 4.3. [17] (Schur's Theorem) *Every square matrix A is unitarily similar to an upper triangular matrix.*

Theorem 4.4. [14] *For a square matrix A,*

$$tr(A^*A) \geq \sum_{\lambda \in \sigma(A)} |\lambda|^2.$$

Proof. From Schur's theorem it follows that

$$U^*AU = T \tag{4.11}$$

where T is upper triangular and U is unitary.

From (4.11) we obtain

$$U^*A^*U = T^* \tag{4.12}$$

Multiplying (4.11) and (4.12) gives

$$U^*A^*AU = T^*T \tag{4.13}$$

Taking the trace of (4.13)

$$tr(T^*T) = tr(U^*A^*AU) \tag{4.14}$$

$$= tr(UU^*A^*A) \tag{4.15}$$

$$= tr(A^*A) \tag{4.16}$$

Using Theorem 4.2 in (4.16) yields

$$tr(T^*T) = \sum_i \sum_k |t_{ki}|^2 \tag{4.17}$$

$$= \sum_k |t_{kk}|^2 + \sum_{i>k} |t_{ki}|^2 \tag{4.18}$$

$$= \sum_{\lambda \in \sigma(A)} |\lambda|^2 + \sum_{i>k} |t_{ki}|^2 \tag{4.19}$$

$$\geq \sum_{\lambda \in \sigma(A)} |\lambda|^2 \tag{4.20}$$

□

Theorem 4.5. [22] *Let A be an $n \times n$ complex matrix with real eigenvalues λ_j such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, and let*

$$m = \frac{tr(A)}{n}, \tag{4.21}$$

$$s^2 = \frac{tr(A^2)}{n} - m^2 \tag{4.22}$$

Then

$$m - s\sqrt{n-1} \leq \lambda_n \leq m - \frac{s}{\sqrt{n-1}} \tag{4.23}$$

$$m + \frac{s}{\sqrt{n-1}} \leq \lambda_1 \leq m + s\sqrt{n-1} \tag{4.24}$$

The equality holds on the left (right) of (4.23) *if and only if equality holds on the left (right) of* (4.24) *if and only if the* $n-1$ *largest (smallest) eigenvalues are equal.*

Firstly we require to prove the following two lemmas in order to establish Theorem 4.5.

Lemma 4.6. [22] *Let* $\mathbf{w}$ *and* $\boldsymbol{\lambda}$ *be real non-zero* $n \times 1$ *vectors and let*

$$m = \frac{\boldsymbol{\lambda}^T \mathbf{e}}{n} \tag{4.25}$$

and

$$s^2 = \frac{\langle C\boldsymbol{\lambda}, \boldsymbol{\lambda} \rangle}{n} \tag{4.26}$$

where $\mathbf{e}$ *is the* $n \times 1$ *vector of ones,* $C = I - \mathbf{e}\mathbf{e}^T/n$, *and* $\mathbf{e}^T$ *is the transpose of* $\mathbf{e}$. *Then*

$$-s\sqrt{n\mathbf{w}^T C\mathbf{w}} \leq \mathbf{w}^T\boldsymbol{\lambda} - m\mathbf{w}^T\mathbf{e} = \mathbf{w}^T C\boldsymbol{\lambda} \leq s\sqrt{n\mathbf{w}^T C\mathbf{w}} \tag{4.27}$$

Equality holds on the left (right) of (4.27) *if and only if*

$$\boldsymbol{\lambda} = a\mathbf{w} + b\mathbf{e} \tag{4.28}$$

for some scalars a *and* b, *where* $a < 0$ $(a > 0)$.

Proof. We observe that $\mathbf{e}\mathbf{e}^T = [\mathbf{e}\ \mathbf{e}\ \cdots\ \mathbf{e}]$, hence $\mathrm{rank}\big(\mathbf{e}\mathbf{e}^T\big) = 1$. This implies that C is rank deficient.

We now establish the equivalence of (4.22) and (4.26).

$$\frac{\langle C\boldsymbol{\lambda}, \boldsymbol{\lambda} \rangle}{n} = \frac{1}{n}\left\langle \left(I - \frac{\mathbf{e}\mathbf{e}^T}{n}\right)\boldsymbol{\lambda}, \boldsymbol{\lambda} \right\rangle \tag{4.29}$$

$$= \frac{1}{n}\left(\|\boldsymbol{\lambda}\|^2 - \frac{\mathbf{e}^T\boldsymbol{\lambda}}{n}\langle \mathbf{e}, \boldsymbol{\lambda} \rangle\right) \tag{4.30}$$

$$= \frac{1}{n}\left(\|\boldsymbol{\lambda}\|^2 - \frac{\langle \mathbf{e}, \boldsymbol{\lambda} \rangle^2}{n}\right) \tag{4.31}$$

$$= \frac{tr(A^2)}{n} - \left(\frac{tr(A)}{n}\right)^2 \tag{4.32}$$

$$= \frac{tr(A^2)}{n} - m^2. \tag{4.33}$$

We now show that C is idempotent.

$$C^2 = \left(I - \frac{\mathbf{e}\mathbf{e}^T}{n}\right)\left(I - \frac{\mathbf{e}\mathbf{e}^T}{n}\right) \tag{4.34}$$

$$= I - \frac{2\mathbf{e}\mathbf{e}^T}{n} + \frac{\mathbf{e}(\mathbf{e}^T\mathbf{e})\mathbf{e}^T}{n^2} \tag{4.35}$$

$$= I - \frac{2\mathbf{e}\mathbf{e}^T}{n} + \frac{\mathbf{e}\mathbf{e}^T}{n} \tag{4.36}$$

$$= I - \frac{\mathbf{e}\mathbf{e}^T}{n} \tag{4.37}$$

$$= C \tag{4.38}$$

We now establish the inequality in (4.27). Using the fact that C is idempotent and Hermitian we get

$$\mathbf{w}^T C\boldsymbol{\lambda} = \langle C\boldsymbol{\lambda}, \mathbf{w}\rangle \tag{4.39}$$

$$= \langle C^2\boldsymbol{\lambda}, \mathbf{w}\rangle \tag{4.40}$$

$$= \langle C\boldsymbol{\lambda}, C\mathbf{w}\rangle \tag{4.41}$$

The Cauchy Schwarz inequality applied to (4.41) yields

$$|\mathbf{w}^T C\boldsymbol{\lambda}| = |\langle C\boldsymbol{\lambda}, C\mathbf{w}\rangle| \tag{4.42}$$

$$\leq \sqrt{\langle C\boldsymbol{\lambda}, C\boldsymbol{\lambda}\rangle\langle C\mathbf{w}, C\mathbf{w}\rangle} \tag{4.43}$$

$$= \sqrt{\langle C\boldsymbol{\lambda}, \boldsymbol{\lambda}\rangle\langle C\mathbf{w}, \mathbf{w}\rangle} \tag{4.44}$$

Using (4.26) in (4.44),we obtain

$$|\mathbf{w}^T C\boldsymbol{\lambda}| \leq \sqrt{\langle C\boldsymbol{\lambda}, \boldsymbol{\lambda}\rangle}\sqrt{\langle C\mathbf{w}, \mathbf{w}\rangle} \tag{4.45}$$

$$\leq s\sqrt{n}\sqrt{\langle C\mathbf{w}, \mathbf{w}\rangle} \tag{4.46}$$

$$= s\sqrt{n\mathbf{w}^T C\mathbf{w}}. \tag{4.47}$$

We now prove the equality in (4.27)

$$
\begin{aligned}
\mathbf{w}^T\boldsymbol{\lambda} - m\mathbf{w}^T\mathbf{e} &= \mathbf{w}^T(\boldsymbol{\lambda} - m\mathbf{e}) && (4.48)\\
&= \mathbf{w}^T\left(\boldsymbol{\lambda} - \frac{(\boldsymbol{\lambda}^T\mathbf{e})\mathbf{e}}{n}\right) && (4.49)\\
&= \mathbf{w}^T\left(\boldsymbol{\lambda} - \frac{(\mathbf{e}^T\boldsymbol{\lambda})\mathbf{e}}{n}\right) && (4.50)\\
&= \mathbf{w}^T\left(\boldsymbol{\lambda} - \frac{\mathbf{e}\mathbf{e}^T\boldsymbol{\lambda}}{n}\right) && (4.51)\\
&= \mathbf{w}^T\left(I - \frac{\mathbf{e}\mathbf{e}^T}{n}\right)\boldsymbol{\lambda} && (4.52)\\
&= \mathbf{w}^TC\boldsymbol{\lambda}. && (4.53)
\end{aligned}
$$

Consider the equation

$$C\boldsymbol{\lambda} = aC\mathbf{w} \tag{4.54}$$

We now show the equivalence of (4.28) and (4.54). It follows from the definition of C that $C\mathbf{e} = \mathbf{0}$. From (4.54) we get

$$C(\boldsymbol{\lambda} - a\mathbf{w}) = \mathbf{0}. \tag{4.55}$$

Hence $\boldsymbol{\lambda} - a\mathbf{w}$ belongs to the null space of C, but C has rank $n-1$, hence the nullity of C is 1. Now $C\mathbf{e} = \mathbf{0}$ implies that $\{\mathbf{e}\}$ is a basis for the null space of C, from which it follows that $\boldsymbol{\lambda} - a\mathbf{w} = b\mathbf{e}$ for same scalar b. $\square$

Lemma 4.7. [22] *Let $\boldsymbol{\lambda} = (\lambda_j)$, m and s be defined as in Lemma 4.6, and*

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \tag{4.56}$$

then (*refer to* (4.23) *and* (4.24))

$$\lambda_n \leq m - \frac{s}{\sqrt{n-1}} \leq m + \frac{s}{\sqrt{n-1}} \leq \lambda_1 \tag{4.57}$$

The equality holds on the left if and only if $\lambda_2 = \lambda_3 = \cdots = \lambda_n$, on the right if and only if $\lambda_1 = \lambda_2 = \cdots = \lambda_{n-1}$, and in the center if and only if $\lambda_1 = \lambda_2 = \cdots = \lambda_n$, if and only if $s = 0$.

Proof. We have that

$$
\begin{aligned}
n^2(m-\lambda_n)^2 &= n^2\left(\frac{\sum_{i=1}^{n}\lambda_i}{n} - \lambda_n\right)^2 && (4.58)\\
&= n^2\left(\frac{\sum_{i=1}^{n}\lambda_i - n\lambda_n}{n}\right)^2 && (4.59)\\
&= \left(\sum_{i=1}^{n}\lambda_i - n\lambda_n\right)^2 && (4.60)\\
&= \left[\sum_{i=1}^{n}(\lambda_i - \lambda_n)\right]^2 && (4.61)\\
&= \sum_{i=1}^{n}(\lambda_i-\lambda_n)^2 + \sum_{i\neq k}(\lambda_i-\lambda_n)(\lambda_k-\lambda_n) && (4.62)\\
&\geq \sum_{i=1}^{n}(\lambda_i-\lambda_n)^2 && (4.63)\\
&= \langle \boldsymbol{\lambda} - \lambda_n\mathbf{e}, \boldsymbol{\lambda} - \lambda_n\mathbf{e}\rangle && (4.64)\\
&= \langle \boldsymbol{\lambda} - m\mathbf{e} + m\mathbf{e} - \lambda_n\mathbf{e}, \boldsymbol{\lambda} - m\mathbf{e} + m\mathbf{e} - \lambda_n\mathbf{e}\rangle && (4.65)\\
&= \langle \boldsymbol{\lambda} - m\mathbf{e}, \boldsymbol{\lambda} - m\mathbf{e}\rangle + \langle m\mathbf{e} - \lambda_n\mathbf{e}, m\mathbf{e} - \lambda_n\mathbf{e}\rangle \\
&+ 2\langle \boldsymbol{\lambda} - m\mathbf{e}, m\mathbf{e} - \lambda_n\mathbf{e}\rangle && (4.66)\\
&= \langle C\boldsymbol{\lambda}, C\boldsymbol{\lambda}\rangle + (m-\lambda_n)^2\langle \mathbf{e}, \mathbf{e}\rangle + 2\langle C\boldsymbol{\lambda}, (m-\lambda_n)\mathbf{e}\rangle && (4.67)\\
&= \langle C\boldsymbol{\lambda}, \boldsymbol{\lambda}\rangle + (m-\lambda_n)^2 n + 2(m-\lambda_n)\langle \boldsymbol{\lambda}, C\mathbf{e}\rangle && (4.68)\\
&= ns^2 + (m-\lambda_n)^2 n && (4.69)
\end{aligned}
$$

We now prove an inequality similar to (4.69)

$$n^2(\lambda_1 - m)^2 = n^2 \left(\lambda_1 - \frac{\sum_{i=1}^{n} \lambda_i}{n} \right)^2 \tag{4.70}$$

$$= \left(n\lambda_1 - \sum_{i=1}^{n} \lambda_i \right)^2 \tag{4.71}$$

$$= \left[\sum_{i=1}^{n} (\lambda_1 - \lambda_i) \right]^2 \tag{4.72}$$

$$= \sum_{i=1}^{n} (\lambda_1 - \lambda_i)^2 + \sum_{i \neq k} (\lambda_1 - \lambda_i)(\lambda_1 - \lambda_k) \tag{4.73}$$

$$\geq \sum_{i=1}^{n} (\lambda_1 - \lambda_i)^2 \tag{4.74}$$

$$= \langle \lambda_1 \mathbf{e} - \boldsymbol{\lambda}, \lambda_1 \mathbf{e} - \boldsymbol{\lambda} \rangle \tag{4.75}$$

$$= \langle \lambda_1 \mathbf{e} - m\mathbf{e} + m\mathbf{e} - \boldsymbol{\lambda}, \lambda_1 \mathbf{e} - m\mathbf{e} + m\mathbf{e} - \boldsymbol{\lambda} \rangle \tag{4.76}$$

$$= \langle \lambda_1 \mathbf{e} - m\mathbf{e}, \lambda_1 - m\mathbf{e} \rangle + \langle m\mathbf{e} - \boldsymbol{\lambda}, m\mathbf{e} - \boldsymbol{\lambda} \rangle$$

$$+ 2\langle \lambda_1 \mathbf{e} - m\mathbf{e}, m\mathbf{e} - \boldsymbol{\lambda} \rangle$$

$$= (\lambda_1 - m)^2 \langle \mathbf{e}, \mathbf{e} \rangle + \langle -C\boldsymbol{\lambda}, -C\boldsymbol{\lambda} \rangle + 2\langle (\lambda_1 - m)\mathbf{e}, -C\boldsymbol{\lambda} \rangle \tag{4.77}$$

$$= (\lambda_1 - m)^2 n + \langle C\boldsymbol{\lambda}, \boldsymbol{\lambda} \rangle - 2(\lambda_1 - m)\langle C\mathbf{e}, \boldsymbol{\lambda} \rangle \tag{4.78}$$

$$= (\lambda_1 - m)^2 n + ns^2 \tag{4.79}$$

From inequality (4.69)

$$n(m - \lambda_n)^2 \geq s^2 + (m - \lambda_n)^2 \tag{4.80}$$

$$\frac{(n-1)(m - \lambda_n)^2}{n-1} \geq \frac{s^2}{n-1} \tag{4.81}$$

$$(m - \lambda_n)^2 \geq \frac{s^2}{n-1} \tag{4.82}$$

$$(m - \lambda_n) \geq \frac{s}{\sqrt{n-1}} \tag{4.83}$$

$$m - \frac{s}{\sqrt{n-1}} \geq \lambda_n \tag{4.84}$$

From inequality (4.79)

$$(\lambda_1 - m)^2 \geq \frac{s^2}{n} + \frac{(\lambda_1 - m)^2}{n} \tag{4.85}$$

$$(\lambda_1 - m)^2 \left(1 - \frac{1}{n}\right) \geq \frac{s^2}{n} \tag{4.86}$$

$$(\lambda_1 - m)^2 \geq \left(\frac{n}{n-1}\right)\left(\frac{s^2}{n}\right) \tag{4.87}$$

$$(\lambda_1 - m)^2 \geq \frac{s^2}{n-1} \tag{4.88}$$

$$(\lambda_1 - m) \geq \frac{s}{\sqrt{n-1}} \tag{4.89}$$

$$\lambda_1 \geq m + \frac{s}{\sqrt{n-1}} \tag{4.90}$$

Equality holding on the left of (4.57) is equivalent to

$$\sum_{i \neq k} (\lambda_i - \lambda_n)(\lambda_k - \lambda_n) = 0 \tag{4.91}$$

from (4.62). Since this is the sum of positive quantities, it implies that $\lambda_2 = \lambda_3 = \cdots = \lambda_n$. If $\lambda_1 = \lambda_n$, then all $\lambda_i\ 's$ are equal which is the trivial case. If $\lambda_2 = \lambda_3 = \cdots = \lambda_n$ then (4.91) is trivially true. Similarly the equality on the right of (4.57) can be shown.

Equality holding in the centre of (4.57) is equivalant to $s = 0$, which from (4.26) implies that $\langle\, C\boldsymbol{\lambda}, \boldsymbol{\lambda}\rangle = 0$. Since C is symmetric and idempotent it follows that $\langle\, C\boldsymbol{\lambda}, C\boldsymbol{\lambda}\rangle = 0$, which implies that $C\boldsymbol{\lambda} = \mathbf{0}$. From the definition of C it is easy to show that $\boldsymbol{\lambda} = m\mathbf{e}$. Thus all the lambdas are equal.

Now substituting $\mathbf{w} = \mathbf{e}_j$ in (4.27) we simplify the right hand side to get

$$s\sqrt{n\mathbf{w}^T C\mathbf{w}} = s\sqrt{n\mathbf{e}_j^T C\mathbf{e}_j} \tag{4.92}$$

$$= s\sqrt{n\mathbf{e}_j^T \left(I - \frac{\mathbf{e}\mathbf{e}^T}{n}\right)\mathbf{e}_j} \tag{4.93}$$

$$= s\sqrt{n\mathbf{e}_j^T\mathbf{e}_j - \mathbf{e}_j^T\mathbf{e}\mathbf{e}^T\mathbf{e}_j} \tag{4.94}$$

$$= s\sqrt{n-1} \tag{4.95}$$

Thus (4.27) becomes

$$-s\sqrt{n-1} \le \lambda_j - m \le s\sqrt{n-1} \tag{4.96}$$

This proves the left hand side of (4.23) and the right hand side of (4.24) by choosing $j = n$ and $j = 1$ in (4.96).

Now we assume that equality holds in the left of (4.23), then

$$m - s\sqrt{n-1} = \lambda_n \tag{4.97}$$

$$m - \lambda_n = s\sqrt{n-1} \tag{4.98}$$

$$\frac{m}{n-1} - \frac{\lambda_n}{n-1} = \frac{s\sqrt{n-1}}{n-1} \tag{4.99}$$

$$m + \frac{m}{n-1} - \frac{\lambda_n}{n-1} = m + \frac{s}{\sqrt{n-1}} \tag{4.100}$$

$$\frac{m(n-1) + m - \lambda_n}{n-1} = m + \frac{s}{\sqrt{n-1}} \tag{4.101}$$

$$\frac{mn - \lambda_n}{n-1} = m + \frac{s}{\sqrt{n-1}} \tag{4.102}$$

$$\frac{\lambda_1 + \lambda_2 + \cdots + \lambda_{n-1}}{n-1} = m + \frac{s}{\sqrt{n-1}} \tag{4.103}$$

If $\lambda_1 = \lambda_2 = \cdots = \lambda_{n-1}$, that is the $n-1$ largest eigenvalues are equal, then from

(4.103) we get

$$\lambda_1 = m + \frac{s}{\sqrt{n-1}} \tag{4.104}$$

If (4.104) is true, the $n-1$ largest eigenvalues are equal, then from (4.104)

$$\lambda_1(n-1) = m(n-1) + s\sqrt{n-1} \tag{4.105}$$

$$= mn - m + s\sqrt{n-1} \tag{4.106}$$

$$= (n-1)\lambda_1 + \lambda_n - m + s\sqrt{n-1} \tag{4.107}$$

from which (4.97) follows.

Now we assume that equality holds on the right hand side of (4.24), then

$$m + s\sqrt{n-1} = \lambda_1 \tag{4.108}$$

$$m - \lambda_1 = -s\sqrt{n-1} \tag{4.109}$$

$$\frac{m-\lambda_1}{n-1} = -\frac{s}{\sqrt{n-1}} \tag{4.110}$$

$$m + \frac{m-\lambda_1}{n-1} = m - \frac{s}{\sqrt{n-1}} \tag{4.111}$$

$$\frac{m(n-1)+m-\lambda_1}{n-1} = m - \frac{s}{\sqrt{n-1}} \tag{4.112}$$

$$\frac{mn-\lambda_1}{n-1} = m - \frac{s}{\sqrt{n-1}} \tag{4.113}$$

$$\frac{\lambda_2+\lambda_3+\cdots+\lambda_n}{n-1} = m - \frac{s}{\sqrt{n-1}} \tag{4.114}$$

If $\lambda_2 = \lambda_3 = \cdots = \lambda_n$, that is the $n-1$ smallest eigenvalues are equal, then from (4.114) we get

$$\lambda_n = m - \frac{s}{\sqrt{n-1}} \tag{4.115}$$

If (4.115) is true, the $n-1$ smallest eigenvalues are equal, then from (4.115)

$$\lambda_n(n-1) = m(n-1) - s\sqrt{n-1} \tag{4.116}$$

$$= mn - m - s\sqrt{n-1} \tag{4.117}$$

$$= (n-1)\lambda_n + \lambda_1 - m - s\sqrt{n-1} \tag{4.118}$$

from which (4.108) follows.

We observed that when $n = 2$, the two inequalities (4.23) and (4.24) collapse to yield $\lambda_n = m - s$ and $\lambda_1 = m + s$. Lemma 4.7 proves the right hand side of (4.23) and the left hand side of (4.24), the theorem is established. □

4.1 Examples

Example 7. Consider the matrix A defined by

$$A = \begin{bmatrix} 7.4918 + 6.5902i & 0.7869 + 3.3443i & 4.9836 + 2.1803i & 5.5902 + 5.5082i \\ -7.1148 - 5.7377i & 2.2164 - 3.5803i & -5.8295 + 0.7246i & -7.5377 - 0.2852i \\ 5.2131 + 2.6557i & 1.1410 + 1.8492i & 7.6262 - 2.0885i & 5.2557 - 3.4131i \\ -1.2787 - 3.9344i & 0.3541 - 1.4951i & -2.3572 - 0.2689i & 0.6656 - 0.9213i \end{bmatrix}$$

which has real eigenvalues 1, 4, 5 and 8. The bounds for λ_n given by (4.23) are $0.1699 \leq \lambda_n \leq 3.0566$, and for λ_1 given by (4.24) are $5.9434 \leq \lambda_1 \leq 8.8301$.

If $\lambda = x + iy$ and the Gershgorin circle is centred at $a + ib$ with radius r, then

$$|(x-a) + i(y-b)| \leq r \tag{4.119}$$

which implies that the real eigenvalues ($y = 0$) lie in the union of the intervals $[x_L, x_R]$, where $x_L = a - \sqrt{r^2 - b^2}$ and $x_R = a + \sqrt{r^2 - b^2}$. These values are represented in Table 4.1.

Table 4.1

Centre	x_L	x_R
A(1,1)	-7.8781	22.8617
A(2,2)	-20.6235	25.0563
A(3,3)	-6.5105	21.7629
A(4,4)	-7.3276	8.6588

Hence the Gershgorin theorem implies that $\lambda_n \geq -20.6235$ and $\lambda_1 \leq 25.0563$.

Obviously Theorem 4.5 gives superior results.

Chapter 5

Special Tridiagonal Matrices

In this chapter we consider bounds related to tridiagonal matrices J of the form

$$J = \begin{bmatrix} \alpha & \beta & & & & \\ \beta & \alpha & \gamma & & & \\ & \gamma & \alpha & \ddots & & \\ & & \ddots & \ddots & \gamma & \\ & & & \gamma & \alpha & \delta \\ & & & & \delta & \alpha \end{bmatrix} \in \mathbb{R}^{m \times m} \tag{5.1}$$

with given entries $\alpha, \beta, \gamma, \delta \in \mathbb{R}$.

Firstly we consider matrices of the form

$$T = \begin{bmatrix} \alpha & \beta_2 & & & \\ \beta_2 & \alpha & \beta_3 & & \\ & \beta_3 & \alpha & \ddots & \\ & & \ddots & \ddots & \beta_m \\ & & & \beta_m & \alpha \end{bmatrix} \tag{5.2}$$

and

$$T_0 = \begin{bmatrix} 0 & \beta_2 & & & \\ \beta_2 & 0 & \beta_3 & & \\ & \beta_3 & 0 & \ddots & \\ & & \ddots & \ddots & \beta_m \\ & & & \beta_m & 0 \end{bmatrix}$$

where we assume that $\beta_j \neq 0 \ \ \forall j = 2, \ldots, m$. Before proving the main result, we prove the following statements.

Lemma 5.1. [4] *We have $\sigma(T) = \alpha + \sigma(T_0)$, and the spectrum of T_0 is symmetric around the origin.*

Proof. The first observation follows immediately from $T = \alpha I + T_0$, whereas the second statement comes from the fact that T_0 is similar to $-T_0$ by a simple diagonal similarity transformation using $D = \text{diag}[+1, -1, +1, -1, \cdots, (-1)^{m+1}]$ given by.

$$DT_0D = -T_0 \tag{5.3}$$

Hence T_0 and $-T_0$ have the same spectrum. If λ is an eigenvalue of T_0, then λ is also an eigenvalue of $-T_0$, which implies $-\lambda$ is eigenvalue of T_0.

If we have a good upper bound K for $\lambda_1(T_0)$, then $-K$ is the corresponding lower bound for $\lambda_n(T_0)$, that is, we have $\sigma(T_0) \subseteq [-K, +K]$ or, equivalently, $\sigma(T) \subseteq [\alpha - K, \alpha + K]$. Now, the task is to obtain suitable bounds $K > 0$. □

Proposition 5.1.1. [4] *Given $K > 0$, let the sequence $y_1, \ldots, y_m$ be defined by*

$$y_1 = K, \qquad y_{j+1} = K - \frac{\beta_{j+1}^2}{y_j}, \qquad \forall j = 1, \ldots, m-1 \tag{5.4}$$

and suppose that $y_j > 0 \ \ \forall j = 2, \ldots, m-2$ *and* $y_{m-1} \geq \frac{\beta_m^2}{K}$*. Then the spectrum of* T_0 *is contained in the interval* $[-K, +K]$*.*

Proof. The assumptions imply that $y_j > 0$ for all $j = 1, \ldots, m-1$ and $y_m \geq 0$. Hence the matrix

$$L = \begin{bmatrix} \sqrt{y_1} & & & & \\ b_2 & \sqrt{y_2} & & & \\ & b_3 & \sqrt{y_3} & & \\ & & \ddots & \ddots & \\ & & & b_m & \sqrt{y_m} \end{bmatrix}, \quad b_j = -\frac{\beta_j}{\sqrt{y_{j-1}}}, \forall_j = 2, \ldots, m \tag{5.5}$$

is well defined, and by calculations we can show that we have the Cholesky-decomposition [5] $KI - T_0 = LL^T$. Hence $KI - T_0$ is positive semi-definite. From this it follows that $\sigma(KI - T_0) \geq 0$ which implies that $\sigma(T_0) \leq K$. Since the spectrum of T_0 is symmetric with respect to the origin according to Lemma 5.1, it follows that $\sigma(KI - T_0) \subseteq [0, 2K]$. □

Lemma 5.2. [4] *We have that* $\sigma(J) \subseteq [\alpha - K, \alpha + K]$ *with* $K \geq 2|\gamma|$

Proof. Since J is a special form of T, $\sigma(J) \subseteq [\alpha - K, \alpha + K]$ follows from Lemma 5.1. We consider the permutation matrix P defined by

$$P_{i,i+1} = 1 \qquad i = 1, 2, \cdots, m-1$$

$$P_{m,1} = 1.$$

Then PJP^T has the form

$$PJP^T = \left[\begin{array}{cccccc|cc}
\alpha & \gamma & & & & & & \beta \\
\gamma & \alpha & \gamma & & & & & \\
& \gamma & \alpha & \ddots & & & & \\
& & \ddots & \ddots & \gamma & & & \\
& & & \gamma & \alpha & \gamma & & \\
& & & & \gamma & \alpha & \delta & \\
\hline
& & & & & \delta & \alpha & \\
\beta & & & & & & & \alpha
\end{array}\right] \tag{5.6}$$

Now consider the principal sub-matrix $\hat{J}$ of order $(m-2)\times(m-2)$ demarcated in (5.6). We now attempt to find the eigenvalues of $\hat{J}$. Let the eigenvector $\mathbf{x} = [x_1, x_2, \cdots, x_{m-2}]^T$ of $\hat{J}$ correspond to eigenvalue $\hat{\lambda}$, then $\left(\hat{J} - \hat{\lambda} I\right)\mathbf{x} = \mathbf{0}$. The components of $\mathbf{x}$ satisfy the second order difference equation.

$$\gamma x_{k-1} + (\alpha - \hat{\lambda})x_k + \gamma x_{k+1} = 0 \qquad k = 1, 2, \cdots, m-2 \tag{5.7}$$

with $x_0 = x_{m-1} = 0$. Let $x_k = t^k$ then by substituting in (5.7) we get

$$t^{k-1}[\gamma + (\alpha - \hat{\lambda})t + \gamma t^2] = 0. \tag{5.8}$$

Hence

$$\gamma + (\alpha - \hat{\lambda})t + \gamma t^2 = 0 \tag{5.9}$$

giving roots

$$t_1 = \frac{(\hat{\lambda} - \alpha) + \sqrt{(\hat{\lambda} - \alpha)^2 - 4\gamma^2}}{2\gamma} \tag{5.10}$$

and

$$t_2 = \frac{(\hat{\lambda} - \alpha) - \sqrt{(\hat{\lambda} - \alpha)^2 - 4\gamma^2}}{2\gamma}. \tag{5.11}$$

If $t_1 = t$ then $x_k = (C_1 + kC_2)t^k$ where C_1 and C_2 are constants. Then $x_0 = x_{m-1} = 0$ implies that $x_k = 0$, which contradicts the fact that $\mathbf{x}$ is an eigenvector. Hence the solution has the form

$$x_k = C_1 t_1^k + C_2 t_2^k. \tag{5.12}$$

Let

$$\hat{\lambda} - \alpha = 2|\gamma| \cos\theta \tag{5.13}$$

then $t_1 = \frac{|\gamma|}{\gamma} e^{i\theta}$ and $t_2 = \frac{|\gamma|}{\gamma} e^{-i\theta}$. Hence (5.12) becomes

$$x_k = \tfrac{|\gamma|^k}{\gamma^k} \left(C_1 e^{ik\theta} + C_2 e^{-ik\theta} \right). \tag{5.14}$$

Applying the condition $x_0 = 0$ gives $C_1 = -C_2$ and subsequently using the condition $x_{m-1} = 0$ gives

$$e^{i(m-1)\theta} - e^{-i(m-1)\theta} = 0 \tag{5.15}$$

from which it follows that $e^{i2(m-1)\theta} = 1 = e^{i2\pi q}$ for integer values of q. Hence

$$\theta_q = \frac{\pi q}{m-1} \tag{5.16}$$

From (5.13) we get

$$\hat{\lambda}_q = \alpha + 2|\gamma| \cos\left(\frac{\pi q}{m-1}\right) \qquad q = 1, 2, \cdots, m-2 \tag{5.17}$$

Hence $\hat{\lambda}_{min} = \alpha + 2|\gamma| \cos\left(\frac{\pi(m-2)}{m-1}\right)$ and $\hat{\lambda}_{max} = \alpha + 2|\gamma| \cos\left(\frac{\pi}{m-1}\right)$.

We note that J and PJP^T have the same eigenvalues. The eigenvalues of $\hat{J}$ interlace the eigenvalues of PJP^T and hence that of J [17]. From $\lambda_m \leq \hat{\lambda}_{min}$ and $\hat{\lambda}_{max} \leq \lambda_1$ which must hold in the limit as $m \to \infty$, we get $\sigma(\hat{J}) \in [\alpha - 2|\gamma|, \alpha + 2|\gamma|]$. Since $\sigma(J) \subseteq [\alpha - K, \alpha + K]$ we must have that $K \geq 2|\gamma|$. □

Assumption 5.3.

(a) *It holds that* $m \geq 4$.

(b) *It holds that* $\beta\gamma\delta \neq 0$.

(c) *The constant* K *always satisfies* $K \geq 2|\gamma|$.

Assumption (a) is clear since otherwise the matrix J is not defined. Assumption (b) can be stated without loss of generality since otherwise the matrix reduces to similar matrices of smaller dimension, whereas assumption (c) is clear in view of Lemma 5.2. In addition, we may assume without loss of generality that $|\beta| \geq |\delta|$ since it is easy to see that J is similar to a matrix which has the same entries as J except that the roles of β and δ are exchanged.

Define J_0 to be the matrix arising from J by setting all diagonal elements to zero. In view of Lemma 5.1, we know that $\sigma(J) = \alpha + \sigma(J_0)$, and that the eigenvalues of J_0 are symmetrically distributed around the origin. In order to obtain good lower and upper bounds for the extremal eigenvalues of J, it therefore suffices to find a suitable bound $K > 0$ such that $\sigma(J_0) \subseteq [-K, +K]$. We can use Proposition 5.1.1 and the recursion from that result, applied to the matrix J_0, reads as follows:

$$y_1 = K, \quad y_2 = K - \frac{\beta^2}{K}$$

$$y_{j+1} = f(y_j) \ \ \forall j = 2, \cdots, m-2, \ \ \text{where} \ \ f(y) = K - \frac{\gamma^2}{y}, \tag{5.18}$$

$$y_m = K - \frac{\delta^2}{y_{m-1}}.$$

Lemma 5.4. [4] *Let $\gamma \in \mathbb{R}$ and $K > 0$ be given. Choose an initial element $y_1 > 0$ and define $y_{k+1} = f(y_k)$ recursively for $k \in \mathbb{N}$, where f is defined in* (5.18). *Then the following statements holds*

***Case 1*:** *When $K \geq 2|\gamma|$, f has a repelling fixed point $f_1 = \frac{K-\sqrt{K^2-4\gamma^2}}{2}$ and an attracting fixed point $f_2 = \frac{K+\sqrt{K^2-4\gamma^2}}{2}$ which coincide for $K = 2|\gamma|$,that is, $f_1 = f_2$ in this case.*

(a) *For $y_1 \in (f_1, f_2)$ we have $f_1 < y_1 < y_2 < \cdots < y_k < y_{k+1} < \cdots < f_2$ for all $k \in \mathbb{N}$.*

Furthermore, it holds that $\lim_{k\to\infty} y_k = f_2$.

(b) *For $y_1 > f_2$ we have $f_2 < \cdots < y_{k+1} < y_k < \cdots < y_3 < y_2 < y_1$ for all $k \in \mathbb{N}$.*

Furthermore, it holds that $\lim_{k\to\infty} y_k = f_2$

(c) *For $y_1 \in (0, f_1)$ we have $f_1 > y_1 > y_2 > y_3 > \cdots$ and there exists a smallest $k_0 \in \mathbb{N}$ with $y_{k_0} \leq 0$. From that on, the sequence is no longer well-defined.*

Case 2: *When $K < 2|\gamma|$, f has no fixed points.*

Proof. Instead of giving the simple proof, we illustrate this result in Figure 5.1 for (a) only. Similar diagrams can be drawn to illustrate (b) and (c). The fixed points f_1 and f_2 will play an essential role in our analysis; since they depend on the constant K. We will denote them by $f_1(K)$ and $f_2(K)$ from now on. Furthermore, the recursively defined values y_j (where $j = 1, \cdots, m$) also depend on K, so we write $y_j(K)$.

In view of Proposition 5.1.1, we have to find suitable conditions on the matrix entries β, γ, δ of J_0 such that

$$y_j(K) > 0 \;\; \forall j = 2, \cdots, m-2 \;\; \text{and} \;\; y_{m-1}(K) \geq h(K) \tag{5.19}$$

where

$$h(K) = \frac{\delta^2}{K} \tag{5.20}$$

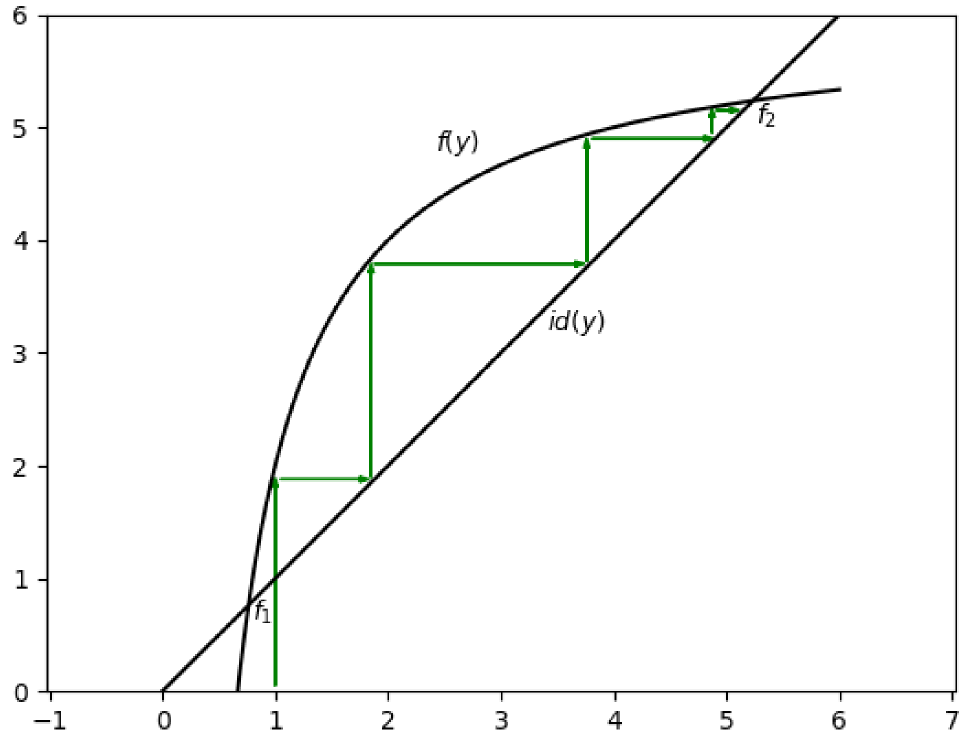


Figure 5.1: Illustration of Lemma 5.4 (a)

□

Lemma 5.5. [4] *Consider the matrix J_0 and assume that $|\beta| > \sqrt{2}|\gamma|$. Then $y_2(K) < f_2(K)$.*

Proof. Firstly we recall from Assumption 5.3 that $K \geq 2|\gamma|$. This implies $1 - \frac{4\gamma^2}{K^2} \in [0, 1)$. Hence, it follows that $\sqrt{1 - \frac{4\gamma^2}{K^2}} \geq 1 - \frac{4\gamma^2}{K^2}$ or, equivalently,

$$K^2 + K\sqrt{K^2 - 4\gamma^2} - 2K^2 + 4\gamma^2 \geq 0. \tag{5.21}$$

Since $2\beta^2 > 4\gamma^2$ by assumption, this yields

$$K^2 + K\sqrt{K^2 - 4\gamma^2} - 2K^2 + 2\beta^2 > 0 \tag{5.22}$$

which may be rewritten as

$$K^2 + K\sqrt{K^2 - 4\gamma^2} > 2K^2 - 2\beta^2. \tag{5.23}$$

Division by $2K$ gives $f_2(K) > y_2(K)$ in view of the definitions of $f_2(K)$ and $y_2(K)$, respectively. □

In the following, we will use the abbreviations

$$\bar{\beta} = \frac{\beta^2}{\sqrt{\beta^2 - \gamma^2}} \quad \text{and} \quad \bar{\delta} = \frac{\delta^2}{\sqrt{\delta^2 - \gamma^2}} \tag{5.24}$$

for $|\beta|, |\delta| > |\gamma|$. Then we have the preliminary result.

Lemma 5.6. [4] *Consider the matrix J_0 and suppose that $|\beta| > \sqrt{2}|\gamma|$. Then the following statements hold:*

(a) $y_2\left(\bar{\beta}\right) = f_1\left(\bar{\beta}\right)$ *and* $y_2(K) > f_1(K)$ *for all* $K > \bar{\beta}$.

(b) *If* $|\delta| \in \left(|\gamma|, \sqrt{2}|\gamma|\right]$, *then* $h(K) < f_2(K)$ *for all* $K > \bar{\beta}$.

(c) *If* $|\delta| > \sqrt{2}|\gamma|$, *then* $h(K) < f_2(K)$ *for all* $K > \bar{\delta}$.

Proof. We begin with some preliminary observations. Let $l, \gamma \in \mathbb{R}$ be given such that $|l| > |\gamma|$, define

$$\bar{l} = \frac{l^2}{\sqrt{l^2 - \gamma^2}}, \tag{5.25}$$

and the strictly increasing function

$$g_l : [2|\gamma|, \infty) \to \mathbb{R}, \qquad g_l(x) = x^2 + x\sqrt{x^2 - 4\gamma^2} - 2l^2. \tag{5.26}$$

Then the following statements hold :

(i) We will always have $\bar{l} \geq 2|\gamma|$, as shown below

$$(l^2 - 2\gamma^2)^2 \geq 0 \tag{5.27}$$

$$l^4 - 4l^2y^2 + 4\gamma^4 \geq 0 \tag{5.28}$$

$$l^4 \geq 4l^2\gamma^2 - 4\gamma^4 \tag{5.29}$$

$$= 4\gamma^2(l^2 - \gamma^2) \tag{5.30}$$

$$\frac{l^4}{l^2 - \gamma^2} \geq 4\gamma^2 \tag{5.31}$$

$$\bar{l} = \frac{l^2}{\sqrt{l^2 - \gamma^2}} \geq 2|\gamma| \tag{5.32}$$

and equality holds if and only if $|l| = \sqrt{2}|\gamma|$.

(ii) If $|l| \leq \sqrt{2}|\gamma|$, then $g_l(x) > 0$ for all $x \in (2|\gamma|, \infty)$, as shown below. Since $x > 2|\gamma|$ we have

$$g_l(x) > g_l\,(2|\gamma|) \tag{5.33}$$

$$= 4\gamma^2 - 2l^2 \geq 0 \tag{5.34}$$

(iii) If $|l| > \sqrt{\gamma}$, then $g_l(\bar{l}) = 0$ and $g_l(x) > 0$ for all $x \in (\bar{l}, \infty)$. This is proved as follows

$$g_l(\bar{l}) = \frac{l^4}{l^2-\gamma^2} + \frac{l^2}{\sqrt{l^2-\gamma^2}}\sqrt{\frac{l^4}{l^2-\gamma^2} - 4\gamma^2} - 2l^2 \tag{5.35}$$

$$= \frac{l^4}{l^2-\gamma^2} + \frac{l^2}{l^2-\gamma^2}\sqrt{(l^2-2\gamma^2)^2} - 2l^2 \tag{5.36}$$

$$= \frac{l^4}{l^2-\gamma^2} + \frac{l^2(l^2-2\gamma^2)}{l^2-\gamma^2} - 2l^2 \tag{5.37}$$

$$= \frac{l^4 + l^2(l^2-2\gamma^2) - 2l^2(l^2-\gamma^2)}{l^2-\gamma^2} = 0. \tag{5.38}$$

Now $x > \bar{l}$ implies

$$g_l(x) > g_l(\bar{l}) = 0. \tag{5.39}$$

(a) By direct computation it can be shown that

$$y_2(\bar{\beta}) = \bar{\beta} - \frac{\beta^2}{\bar{\beta}} \tag{5.40}$$

$$= \frac{\beta^2}{\sqrt{\beta^2-\gamma^2}} - \sqrt{\beta^2-\gamma^2} \tag{5.41}$$

$$= \frac{\beta^2 - (\beta^2-\gamma^2)}{\sqrt{\beta^2-\gamma^2}} \tag{5.42}$$

$$= \frac{\gamma^2}{\sqrt{\beta^2-\gamma^2}} \tag{5.43}$$

$$= f_1(\bar{\beta}). \tag{5.44}$$

Let $l = \beta$, then

$$g_\beta(K) = K^2 + K\sqrt{K^2-4\gamma^2} - 2\beta^2 \tag{5.45}$$

Dividing by $2K$ yields

$$\frac{g_\beta(K)}{2K} = \frac{K}{2} + \frac{\sqrt{K^2-4\gamma^2}}{2} - \frac{\beta^2}{K} \tag{5.46}$$

$$= K - \frac{K}{2} + \frac{\sqrt{K^2-4\gamma^2}}{2} - \frac{\beta^2}{K} \tag{5.47}$$

$$= K - \frac{\beta^2}{K} - \left(\frac{K-\sqrt{K^2-4\gamma^2}}{2}\right) \tag{5.48}$$

$$= y_2(K) - f_1(K). \tag{5.49}$$

Hence the expression $y_2(K) - f_1(K)$ has the same sign as $g_\beta(K)$. Then $K > \bar{\beta} = \bar{l}$ implies from (5.39) that $g_\beta(K) > 0$ which shows that $y_2(K) > f_1(K)$.

For (b), (c) Let $l = \delta$, then

$$g_\delta(K) = K^2 + K\sqrt{K^2-4\gamma^2} - 2\delta^2 \tag{5.50}$$

Dividing by $2K$ yields

$$\frac{g_\delta(K)}{2K} = \frac{K+\sqrt{K^2-4\gamma^2}}{2} - \frac{\delta^2}{K} \tag{5.51}$$

$$= f_2(K) - h(K) \tag{5.52}$$

Hence the sign of $f_2(K) - h(K)$ is the same as sign of $g_\delta(K)$.

(b) Following a similar argument to that of (5.27) to (5.32) by replacing the variable l by β it can be shown that $\bar{\beta} \geq 2|\gamma|$. Since $K > \bar{\beta}$ we have that $K > 2|\gamma|$, let $x = K$ and $l = \delta$ in (5.33) to obtain $g_\delta(K) > 0$. Hence from (5.52) we have $h(K) < f_2(K)$.

(c) Since $\bar{\delta} > |\delta| > 2|\gamma|$ we have that $K > \bar{\delta}$ implies $K > 2|\gamma|$, then from (5.33) and (5.52) we have $h(K) < f_2(K)$. $\square$

Theorem 5.7. [4] *Let $\bar{\beta}$, $\bar{\delta}$ be defined as in* (5.24). *Then the inequalities*

$$\lambda_m(J) \geq \alpha - K \quad \textit{and} \quad \lambda_1(J) \leq \alpha + K$$

hold for the case $|\delta| > \sqrt{2}|\gamma|$, $|\beta| > \sqrt{2}|\gamma|$ *with* $K = \sqrt{\beta^2 + \gamma^2}$

Proof. We may assume without loss of generality that $|\delta| \leq |\beta|$ then

$$(\beta^2 + \delta^2)\gamma^2 \leq 2\beta^2\gamma^2 \tag{5.53}$$

$$< \delta^2\beta^2 \tag{5.54}$$

Inequality (5.54) follows since $|\delta| > \sqrt{2}|\gamma|$. From (5.54) we have

$$\beta^2\gamma^2 + \delta^2\gamma^2 - \delta^2\beta^2 < 0 \tag{5.55}$$

$$-\beta^2\gamma^2 - \delta^2\gamma^2 + \delta^2\beta^2 > 0 \tag{5.56}$$

$$\beta^4 - \beta^2\gamma^2 - \delta^2\gamma^2 + \delta^2\beta^2 > \beta^4 \tag{5.57}$$

$$\beta^2(\beta^2 - \gamma^2) + \delta(\beta^2 - \gamma^2) > \beta^4 \tag{5.58}$$

$$\beta^2 + \delta^2 > \frac{\beta^4}{\beta^2 - \gamma^2} \tag{5.59}$$

$$\sqrt{\beta^2 + \delta^2} > \frac{\beta^2}{\sqrt{\beta^2 - \gamma^2}} \tag{5.60}$$

$$= \bar{\beta} \tag{5.61}$$

Hence $K > \bar{\beta}$. This implies that $y_2(K) > f_1(K)$ from Lemma 5.6(a).

Since $K = \sqrt{\beta^2 + \delta^2}$ we also have

$$K^2 = \beta^2 + \delta^2 \tag{5.62}$$

$$\frac{\delta^2}{K} = K - \frac{\beta^2}{K} \tag{5.63}$$

$$h(K) = y_2(K). \tag{5.64}$$

Now Lemma 5.4 implies that

$$0 < f_1(K) < y_2(K) < y_j(K) \qquad \forall j = 3, \cdots, m-2 \tag{5.65}$$

and

$$\frac{\beta_m^2}{K} = \frac{\delta^2}{K} = h(K) = y_2(K) < y_{m-1}(K) \tag{5.66}$$

Now applying Proposition 5.1.1 implies $\sigma(J_0) \subseteq [-K, K]$. Hence $\sigma(J) \subseteq [\alpha - K, \alpha + K]$ $\square$

5.1 Example

We test the result of Theorem 5.7 by using the matrix J defined by

$$J = \begin{bmatrix} 4 & 3 & 0 & 0 & 0 \\ 3 & 4 & 1 & 0 & 0 \\ 0 & 1 & 4 & 1 & 0 \\ 0 & 0 & 1 & 4 & 2 \\ 0 & 0 & 0 & 2 & 4 \end{bmatrix} \tag{5.67}$$

which has eigenvalues 0.8074, 1.8074, 4.0, 6.1926 and 7.1926. From Theorem 5.7 we obtain $\sigma(J) \subseteq [0.39, 7.606]$ as compared to the Gershgorin circle theorem which gives $\sigma(J) \subseteq [0, 8]$.

Chapter 6

Conclusion

In this study we have investigated bounds for the minimal and maximal eigenvalues of matrices. Firstly we considered matrices from $\mathbb{C}^{n\times n}$ and bounded them by using the entries of the matrix itself. This rather crude approach gives relatively poor bounds. Better bounds are obtained using the Gershgorin circle theorem as well as the ovals of Cassini. These illustrate the region in complex plane that contains the spectrum of the matrix. It cannot be concluded that one is better then the other, this obviously depends on the matrix.

We then considered partitioning the matrix into square diagonal blocks and establishing bounds by using the spectral norm. This approach can be useful as it may decrease the region in the complex plane that contains the spectrum.

Bounds by traces are used to find intervals containing both the smallest and largest eigenvalues. However this applies only to matrices with real eigenvalues. For locating the smallest and largest eigenvalues, this is a rather powerful technique.

The eigenvalues and eigenvectors of symmetric real tridiagonal matrices with constant super diagonal, main diagonal and sub diagonal are known explicitly. However the discretization of boundary value problems by finite difference techniques destroys this constant structure. For such matrices bounds for the spectrum are found in Chapter 5 which are always superior to Gershgorin bounds.

There are special techniques which can be applied to positive definite matrices. However this is not the focus of this thesis.

Bibliography

[1] Bird J., (2017), *Higher Engineering Mathematics (Eighth Ed.)*, Routledge Taylor and Francies Group, ISBN: 1138673571.

[2] Brauer A., (1947) *Limits for the Characteristic Roots of a Matrix*, II. Duke Math. J., Vol. 14, 21-26.

[3] Bodewig E., (1959), *Matrix Calculus, Second re. Ed.*, Interscience Publ.

[4] Buchholzer H. and Kanzow C., (2011), *Bounds for the Extremal Eigenvalues of a Class of Symmetric Tridiagonal Matrices with Applications*, Linear Algebra and its Appl., doi:10.1016/j.laa.2011.10.013.

[5] Burden R.L. and Faires J.D., (2010), *Numerical Analysis, Ninth Ed.*, Cengage Learning.

[6] Euler M., (1960), *Du mouvement d'un corps solide quelconque lorsqu'il tourne autour d'un axe mobile*, Histoire de l'Acadmie royale des sciences et des belles lettres de Berlin,176227.

[7] Feingold D. G. and Varga R. S., (1962), *Block Diagonally Dominant Matrices and Generalizations of the Gerschgorin Circle Theorem*, Pacific J. Math., vol. 12, 1241-1250.

[8] Francis J. G. F., (1961), *The QR Transformation, I (part 1)*, The Computer Journal, Vol. 4: 265-271, doi:10.1093/comjnl/4.3.265.

[9] Hawkins T., (1975), *Cauchy and the Spectral Theory of Matrices*, Historia Mathematica, Vol. 2, 1-29, doi:10.1016/0315-0860(75)90032-4.

[10] Hirsch M. A. and Bendixson M. I., (Feb. 1902), *Sur les racines D'une Équation Fondamentale*, Acta. Math., vol. 25, 367-370.

[11] Householder A. S., (1955), *On the Convergence of Matrix Iterations*, ORNL-1883 (Contract W-7405-eng-26), U.S. At. Energy Comm.

[12] Kline M., (1972), *Mathematical Thought from Ancient to Modern Times*, Oxford University Press, 1063, ISBN 0-19-501496-0.

[13] Kublanovskaya V. N. (1961), *On some Algorithms for the Solution of the Complete Eigenvalue Problem*, USSR Computational Mathematics and Mathematical Physics, Volume 3, 637-657.

[14] Lancaster P., (1969), *Theory of Matrices*, Academic Press

[15] Langville A. N. and Meyer C. D. (2006), *Google's Pagerank and Beyond: The Science of Search Engine Rankings*, Princeton University Press.

[16] Lohmöller J. B., (1989), *Latent Variable Path Modeling with Partial Least Squares*, Springe-Verlag Berlin Heidelberge, ISBN 979-642-52514-8.

[17] Meyer C.D., (2000), *Matrix Analysis and Applied Linear Algebra*, SIAM.

[18] Ostrowski A., (1951), *Ueber das Nichtverschwinden einer Klasse von Determinanten und die Lokalisierung der charakteristischen Wurzeln von Matrizen*. Compositio Math., vol. 9, 209-226.

[19] Shannon C. E. (July, October 1948), *A Mathematical Theory of Communication*, The Bell System Technical Journal, Volume 27, 379-423, 623-656.

[20] Shinozuka M. and Astill C.J., (1972), *Random Eigenvalue Problems in Structural Analysis*, AIAA Journal, Volume 10, 456-462, doi/abs/10.2514/3.50119.

[21] Weisstein E.W., (2003), *CRC concise Encyclopedia of Mathematics, Second ed*, CRC Press, ISBN 1-58488-347-2.

[22] Wolkowicz H. and Styan G.P.H., *Bounds for Eigenvalues Using Traces*, Linear Algebra and its Appl., doi.org/10.1016/0024-3795(80)90258-X.

[23] Yang L., Xu Z., Østergaard J., Dong Z. Y., Wong K. P. and Ma X., (2011), *Oscillatory Stability and Eigenvalue Sensitivity Analysis of A DFIG Wind Turbine System*, IEEE Transactions on Energy Conversion, 328-339, doi:10.1109/TEC.2010.2091130.